THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[SIXTH SERIES.]

JULY 1922.

I. *The Unsteady Motion produced in a Uniformly Rotating Cylinder of Water by a Sudden Change in the Angular Velocity of the Boundary.* By A. R. McLEOD, *M.A.*, *Fellow of Gonville and Caius College, Cambridge* *.

IN the following paper a comparison is made between the observed angular velocities in a rotating circular cylinder of water, and those calculated by two-dimensional theory in which the effects of the base and the free surface are neglected, and each particle of water is assumed to move in a circle about the axis of rotation. The two cases dealt with are those of unsteady motion in which the cylindrical containing-wall is suddenly started from rest, or suddenly stopped when, with the water, it is rotating uniformly. The measurements are all of surface velocities, because the use of lycopodium particles floating on the surface was necessary, as liquid globules of the same density as the water would not remain at the same depth for any length of time. Three cylinders of approximate diameters 5 cm., 15 cm., and 25 cm. were used, and three angular velocities, viz. 36 r.p.m., 10 r.p.m., and $1\frac{1}{3}$ r.p.m. The small cylinder was four diameters, and the other two cylinders were two diameters long. The observed velocities show a departure from theory, which increases with angular velocity and with the size of the cylinder, but which tends to vanish at very low speeds. The discrepancy is much greater

* Communicated by the Author.

for the "stopping" than for the "starting" experiments. The work was done at the Royal Aircraft Establishment, Farnborough, during the months Feb.–Sept. 1919.

In some later work, not yet published, the discrepancy in the case of the "starting" experiments is traced to the effect of the base; and the theory will probably apply to this motion in very long cylinders, provided eddies do not form owing to initial instability of the water. In the case of the "stopping" experiments, the discrepancy is due to instability and turbulence.

§ 1. *Theoretical.*

In dealing with a rectilinear two-dimensional eddy in an incompressible fluid which contains no sources or sinks, the usual assumption is that particles of the fluid move in circles about the axis of the eddy. This makes the problem one of complete symmetry, and the radius vector r and the time t are the two independent co-ordinates. The equations of motion, when written in cylindrical co-ordinates with the axis of z along the axis of symmetry, reduce to the following forms, in which p denotes pressure, ρ is density, ν is the kinetic viscosity, and $\dot{\phi}$ is the angular velocity about the axis :—

$$\frac{\partial p}{\partial r} = \rho r \dot{\phi}^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

for the pressure, and

$$\frac{\partial^2 \dot{\phi}}{\partial r^2} + \frac{3}{r}\frac{\partial \dot{\phi}}{\partial r} = \frac{1}{\nu}\frac{\partial \dot{\phi}}{\partial t} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

giving the angular velocity. Let us suppose that the angular velocity $\dot{\phi}$ satisfies the conditions

$$\dot{\phi} = \mathrm{F}(r) \text{ for } t = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

$$\dot{\phi} = \phi(t) \quad ,, \quad r = c = \text{radius of cylinder}; \quad . \quad (4)$$

that is, at the initial instant the angular velocity in the cylinder is known to be $\mathrm{F}(r)$ at radius r, while thereafter the rotation of the boundary is prescribed to be $\phi(t)$. The solution of (2) satisfying the conditions (3) and (4) is, in

terms of a series of Bessel functions:

$$\dot{\phi} = \sum_{n=1}^{\infty} \frac{2c\,J_1(\alpha_n r/c)e^{-\nu\alpha_n^2 t/c^2}}{r\{J_0(\alpha_n)\}^2} \int_0^1 \xi^2 F(c\xi) J_1(\alpha_n \xi)\, d\xi$$

$$-\sum_{n=1}^{\infty} \frac{2\nu\alpha_n J_1(\alpha_n r/c)e^{-\nu\alpha_n^2 t/c^2}}{cr J_0(\alpha_n)} \int_0^t \phi(\tau) e^{\nu\alpha_n^2 \tau/c^2} d\tau, \quad . \quad (5)$$

where $J_1(\alpha_n)=0$ and α_n is the nth root of this Bessel function of the first order.

Taking $\phi(t)=\Omega=$constant and $F(r)=0$, we have

$$\frac{\dot{\phi}}{\Omega} = 1 - \frac{2c}{r} \sum_{n=1}^{\infty} \frac{J_1(\alpha_n r/c)e^{-\nu\alpha_n^2 t/c^2}}{\alpha_n J_2(\alpha_n)}. \quad . \quad . \quad . \quad (6)$$

This is the solution for the case in which the water is initially at rest and the cylinder suddenly rotates with a constant angular velocity Ω. The solution is given as a problem in Gray & Mathew's 'Treatise on the Bessel Functions' (Ex. 38, p. 236, 1st edn.).

Taking $\phi(t)=0$ and $F(r)=\Omega=$constant, we find the solution

$$\frac{\dot{\phi}}{\Omega} = \frac{2c}{r} \sum_{n=1}^{\infty} \frac{J_1(\alpha_n r/c)}{\alpha_n J_2(\alpha_n)} e^{-\nu\alpha_n^2 t/c^2}. \quad . \quad . \quad . \quad (7)$$

This is the solution for the case in which the water is initially rotating with a constant angular velocity Ω, and the cylinder is suddenly stopped. The solution is given by Stearne, Q. J. Math. xvii. p. 90 (1881), and Tumlirz, *Sitz. d. k. Akad. in Wien*, lxxxv. (ii.) p. 105 (1882).

The phenomena which (6) and (7) are supposed to represent are at the basis of the formation and dissipation of eddies by viscous action. To take one example, they may be of importance in the theory of the aeroplane compass. The experiments of Part I. were undertaken to test the validity of these equations.

§ 2. *Numerical Solution of Equation* (7), *and Discussion.*

The numerical solution of (7) is given in Table I. for the values of r/c and $\nu t/c^2$ which are there indicated. The values are probably accurate throughout to the fifth significant figure. For small values of the arguments, the value of $\dot{\phi}/\Omega$ sometimes differed from unity only after the sixth figure. The values of $J_1(\alpha_n r/c)$ for arguments greater

Table I.

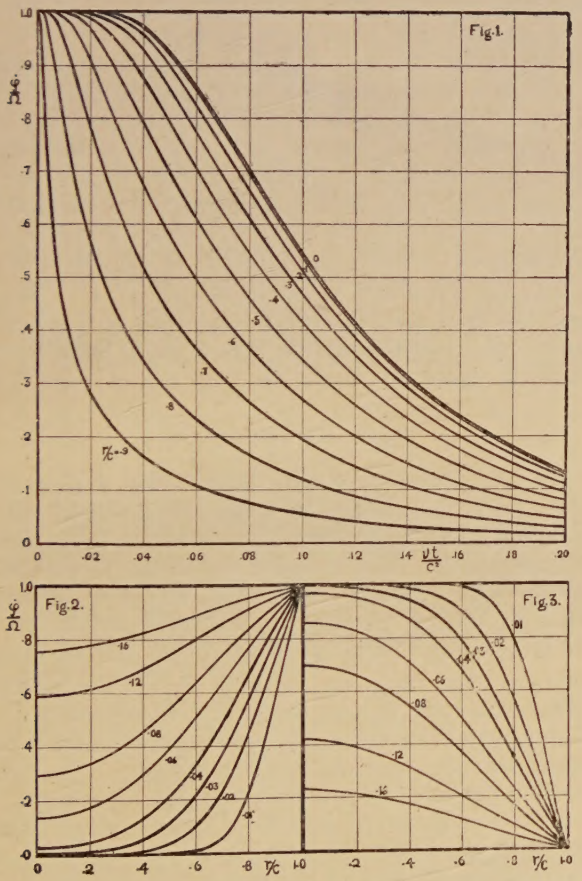
Values of $\dot{\phi}/\Omega$.

vt/c^2.	$r/c = 0$.	·1.	·2.	·3.	·4.
·001	1·00000	1·00000	1·00000	1·00000	1·00000
·002	1·00000	1·00000	1·00000	1·00000	1·00000
·003	1·00000	1·00000	1·00000	1·00000	1·00000
·004	1·00000	1·00000	1·00000	1·00000	1·00000
·006	1·00000	1·00000	1·00000	1·00000	1·00000
·010	1·00000	1·00000	1·00000	·99999	·99992
·015	1·00000	1·00000	·99996	·99969	·99795
·020	·99991	·99982	·99935	·99730	·98970
·025	·99907	·99860	·99650	·99002	·97242
·030	·99587	·99451	·98923	·97591	·94634
·040	·97490	·97062	·95616	·92675	·87504
·050	·92923	·92196	·89890	·85665	·79077
·060	·86299	·85373	·82526	·77594	·70409
·070	·78497	·77482	·74419	·69279	·62086
·080	·70314	·69294	·66248	·61234	·54393
·090	·62300	·61326	·58440	·53746	·47444
·100	·54776	·53877	·51224	·46943	·41257
·150	·27221	·26753	·25321	·23045	·20079
·200	·13156	·12916	·12216	·11104	·09661
·300	·03034	·02979	·02817	·02560	·02226
·400	*·$0^2$6989	·$0^2$6862	·$0^2$6489	·$0^2$5897	·$0^2$5128
·500	·$0^2$1610	·$0^2$1581	·$0^2$1495	·$0^2$1358	·$0^2$1181
·600	·$0^3$3708	·$0^3$3641	·$0^3$3443	·$0^3$3128	·$0^3$2721
·700	·$0^4$8542	·$0^4$8386	·$0^4$7930	·$0^4$7206	·$0^4$6267

vt/c^2.	r/c = ·5.	·6.	·7.	·8.	·9.
·001	1·00000	1·00000	1·00000	·99999	·97034
·002	1·00000	1·00000	1·00000	·99782	·86681
·003	1·00000	1·00000	·99982	·98630	·76998
·004	1·00000	·99998	·99865	·96468	·69193
·006	·99999	·99944	·98952	·90552	·57793
·010	·99886	·99004	·94263	·78057	·44051
·015	·98921	·95570	·85956	·65615	·34305
·020	·96580	·90409	·77540	·56140	·28162
·025	·93062	·84548	·69887	·48810	·23855
·030	·88785	·78592	·63137	·42973	·20630
·040	·79277	·67418	·52033	·34229	·16189
·050	·69768	·57707	·43399	·27956	·12940
·060	·60940	·49443	·36541	·23215	·10650
·070	·53008	·42430	·30983	·19499	·08892
·080	·45994	·36468	·26406	·16509	·07497
·090	·39845	·31382	·22589	·14057	·06365
·100	·34483	·27031	·19375	·12017	·05430
·150	·16624	·12909	·09175	·05653	·02544
·200	·07985	·03190	·04393	·02704	·012155
·300	·018395	·014256	·010114	·$0^2$6223	·$0^2$2797
·400	·$0^2$4237	·$0^2$3284	·$0^2$2330	·$0^2$1433	·$0^3$6443
·500	·$0^3$9760	·$0^3$7563	·$0^3$5366	·$0^3$3301	·$0^3$1484
·600	·$0^3$2248	·$0^3$1742	·$0^3$1236	·$0^4$7605	·$0^4$3418
·700	·$0^4$5178	·$0^4$4013	·$0^4$2847	·$0^4$1752	·$0^5$7874

* ·$0^2$6989 means ·006989.

than 15 were obtained by the use of the first three terms of the asymptotic expansion. For smaller arguments the values were found by interpolation from the twelve-figure tables in Gray & Mathew's 'Treatise.' Curves representing Table I. are shown in fig. 1, from which figs. 2 and 3 have been derived graphically. The calculation was made



with the aid of Chambers's seven-figure logarithms. It was checked throughout, once everywhere and twice or three times in parts.

For large values of t, the angular velocities are proportional to $J_1(\alpha_1 r)/r$ $(0 \leq r \leq 1)$. Most of the experimental curves are roughly of this form for a large value of t.

The solution, being non-dimensional, applies to all sizes

of cylinder, all angular velocities Ω, and all incompressible fluids. The significance of $\nu t/c^2$ is of interest. According to (6) and (7), the behaviour of a cylinder of liquid is exactly similar to that of a cylinder of double the radius in four times the time. Larger eddies should therefore be of relatively longer duration than smaller eddies. If in these same cylinders we have different liquids, the liquid in the larger cylinder must have a viscosity four times that of the liquid in the smaller cylinder, in order that $\dot{\phi}$ may have the same value for the same values of Ω, r/c, and t. For air $\nu = 0{\cdot}14$, and for water $\nu = 0{\cdot}011$ at ordinary temperature, the ratio being about 12. Ignoring the compressibility of air, eddies of the assumed kind should die away much faster in air than in water, the values of c and Ω being the same.

According to (6) and (7), the rates of growth and decay of $\dot{\phi}/\Omega$ in the cylinders are independent of Ω. We shall see from the experiments that this is true only when Ω is very small, or when the radius c is a small fraction of the length of the cylinder.

§ 3. *Experiments.*

In the experiments which have been made, a brass cylinder, bored as accurately as possible and at least two diameters in length, was rotated about a vertical axis at constant speed. The cylinder was filled to within 1 cm. of the top with ordinary tap-water ; and the observations consisted in timing through a measured angle by stop-watch, lycopodium particles floating on the surface, and so deducing the angular velocity $\dot{\phi}$ at the radius r/c selected. The time of the stop-watch observation was also noted on a watch, in order to get the value of t from the instant of starting or stopping the cylinder. To time the lycopodium particles with more accuracy, a horizontal plane glass plate was mounted just over the cylinder. This was marked with ink in circles centring on the axis of rotation. The radii of the circles had the values $r/c = 0{\cdot}3$, $0{\cdot}5$, $0{\cdot}7$, and $0{\cdot}9$, and a fifth circle was added at $r/c = 1{\cdot}0$ to aid in centring the plate over the cylinder. Straight lines, 45° apart, through the centre of the circles served to indicate the angles through which the particles were timed. A large plane mirror, inclined at 45°, was placed on the glass plate to enable observations to be made from the side.

To obtain freedom from vibration, the driving-motor and

reduction gears were mounted on a separate table. The cylinders were clamped against a spinning-table which rotated very freely and easily in ball-bearings. Slight want of truth in the centring of the upper end of the cylinder was corrected by hanging a small weight on the rim, or by means of adjusting screws which worked in the rim of the water-bath and pressed against the cylinder.

Water-baths, providing a 2-inch layer of water around the sides and base, were fitted to each of the cylinders to make the temperature changes less. These rotated with the cylinder. In the case of the larger cylinders especially, it was found that without a bath it was impossible to secure uniform rotation of the water. When the room temperature was rising in the mornings or early afternoons, inward convection currents at the surface carried an excess of angular momentum towards the axis, and the angular velocity at the centre was sometimes for several hours 30 per cent. greater than that of the cylinder itself, the latter being very small. When cooling down, the effect is reversed and the "core" rotates slower than the cylinder. The water-baths made a great improvement, but with the largest cylinder, especially, the water in the bath and in the cylinder had to be mixed thoroughly before each experiment; and the temperature of the water was regulated with an electric heater to be about 0°·1 C. in advance of the room temperature (when rising). Some experiments lasted for nearly an hour, and a difference of 0°·1 C. between the temperatures of room and water was enough to start thermal currents having considerable effects (cumulative) on the observations.

The cleaned surface of the water was lightly dusted with lycopodium, "rafts" not being allowed to form. When the temperature was being adjusted with the electric heater, the water was thoroughly mixed by stirring and bubbling air through it. The mean temperature was obtained for each experiment to 0°·01 C.

At the end of each observation the stop-watch reading (given to 0·1 sec.) was recorded with the angle through which the lycopodium particle had rotated, and the time of the end of the observation as indicated by watch was also written down. All observations in one experiment were made on a selected circle. The cylinder was timed frequently to verify the speed, which was constant to within a per cent. or two, and irregularities were corrected by an adjustable resistance.

The motor was driven from a 200-volt accumulator circuit.

Three sizes of cylinders were used :

Small cylinder.....	$c=c_1=$	2·40 cm.=radius.
Middle „	$c=c_2=$	7·48 „
Large „	$c=c_3=$	12·69 „

Three angular velocities were selected for each cylinder :

Low speed	$\Omega=\Omega_1=$	·1396= $1\frac{1}{3}$ r.p.m.
Middle „	$\Omega=\Omega_2=$	1·047 =10 „
High „	$\Omega=\Omega_3=$	3·770 =36 „

The large and middle cylinders were painted inside. They were bored with an error of a small fraction of a millimetre at the upper rim. The small cylinder was true to 0·05 mm., and its length was four diameters. The others were slightly more than two diameters long. The bases of all were plane and smooth on the inside.

Experiments were made when the water began to rotate from a condition of rest, and also when it was coming to rest from a condition of uniform rotation, the uniformity being ascertained before the cylinder was stopped.

In reducing the observations, the watch-reading was corrected by subtracting half the stop-watch reading to give the mean time t, and the values of $\nu t/c^2$ and $\dot{\phi}/\Omega$ were calculated for each observation. The viscosity of the water, at the mean temperature during the experiment, was taken from Kaye & Laby's Tables. The values of $\dot{\phi}/\Omega$ were plotted on squared paper against the values of $\nu t/c^2$, each observation being represented by a dot. Hence for a selected circle corresponding to the selected value of r/c, for each value of Ω there corresponds a series of dots which lie in a narrow band. This band defines a curve by its median line. Observations were made on the four circles corresponding to $r/c=0{\cdot}3$, $0{\cdot}5$, $0{\cdot}7$, and $0{\cdot}9$ for each of the three cylinders. Thus for each cylinder there were finally eight sheets of curves—four for "starting" and four for "stopping" experiments. Each of these sheets contained three curves corresponding to the three selected velocities, and corresponded to one of the values of r/c. The theoretical curve from Table I. was also added for comparison. Several experiments were made for each set of conditions, to secure a sufficiently dense band of points for each curve. Apart from a number of preliminary

experiments, the results given here are based on 215 distinct experiments.

In taking the stop-watch readings, disturbances of the circular motion of the lycopodium owing to turbulence or eddying were avoided as much as possible, and selections were made of those motions which lay most nearly along the circle considered. This means that when turbulence is increased, the angles through which the particles are timed become less. The effect of the turbulence is merely to broaden the band of points, but the mean motion, represented by the median line, is always well-defined.

§ 4. *Discussion of Results.*

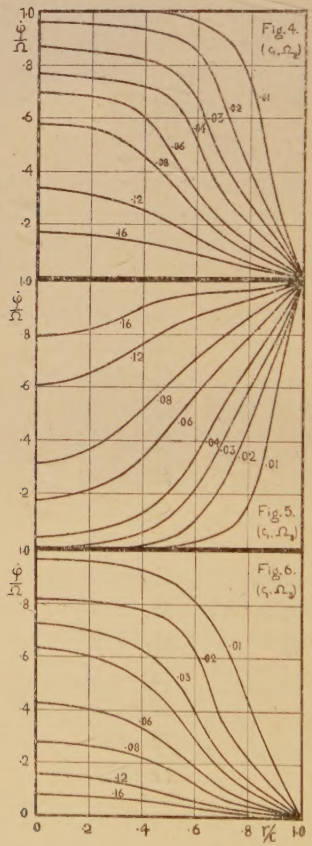
The results obtained are embodied in the accompanying curves. These are derived from the observational curves mentioned in § 3. The following reference table explains the figures :—

Fig. 4.	$c_1 = 2{\cdot}40$ cm.	$\Omega_2 = 10$ r.p.m.	Stopping.
5.	,,	$\Omega_3 = 36$,,	Starting.
6.	,,	$\Omega_3 = 36$,,	Stopping.
7.	$c_2 = 7{\cdot}48$ cm.	$\Omega_1 = 1\frac{1}{3}$,,	Starting.
8.	,,	$\Omega_1 = 1\frac{1}{3}$,,	Stopping.
9.	,,	$\Omega_2 = 10$,,	Starting.
10.	,,	$\Omega_2 = 10$,,	Stopping.
11.	,,	$\Omega_3 = 36$,,	Starting.
12.	,,	$\Omega_3 = 36$,,	Stopping.
13.	$c_3 = 12{\cdot}69$ cm.	$\Omega_1 = 1\frac{1}{3}$,,	Starting.
14.	,,	$\Omega_1 = 1\frac{1}{3}$,,	Stopping.
15.	,,	$\Omega_2 = 10$,,	Starting.
16.	,,	$\Omega_2 = 10$,,	Stopping.
17.	,,	$\Omega_3 = 36$,,	Starting.
18.	,,	$\Omega_3 = 36$,,	Stopping.

In these figures each curve represents the distribution of angular velocities ϕ/Ω over a radius of the cylinder at the time $\nu t/c^2$ corresponding to the number given alongside the curve. The times in seconds corresponding to these numbers may be obtained by use of the following table, the water being at 17° C. :—

$c_1 = 2{\cdot}40$ cm.	$\nu t/c^2 = {\cdot}01$	corresponds to	$t = 5{\cdot}30$	secs.
$c_2 = 7{\cdot}48$,,	$\nu t/c^2 = {\cdot}01$	,,	,,	$t = 51{\cdot}8$	,,
$c_3 = 12{\cdot}69$,,	$\nu t/c^2 = {\cdot}01$	,,	,,	$t = 148{\cdot}4$	,,

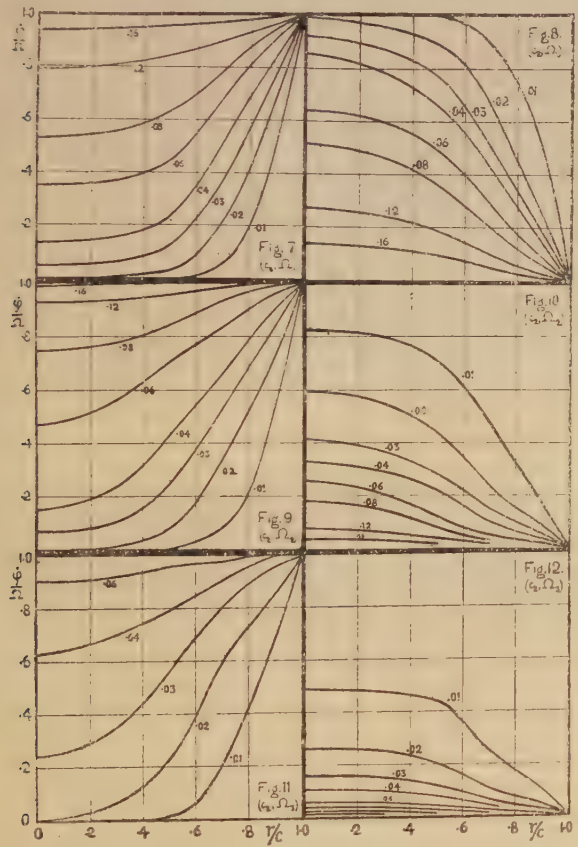
Comparison with the theoretical curves (figs. 2 & 3) shows a marked departure in all cases but that of the slow speed and small cylinder. The agreement with theory improves as $r \rightarrow c$, but there is still a large departure at $r/c = \cdot 9$ for the large and middle cylinders, except at the low speed.



Again, it is noticeable that as the radius of the cylinder increases, the departure from theory becomes more marked. This might be expected as there is, near the axis, relatively less constraint from the boundaries with the larger cylinders. Accordingly, if we suppose an eddy of this kind rotating in a lake of stationary water, and if instead of stopping the cylinder wall we annihilate it, we expect the eddy to disappear more quickly than if the stationary solid wall had

been retained, for the constraint will be still further reduced and greater irregularity is possible.

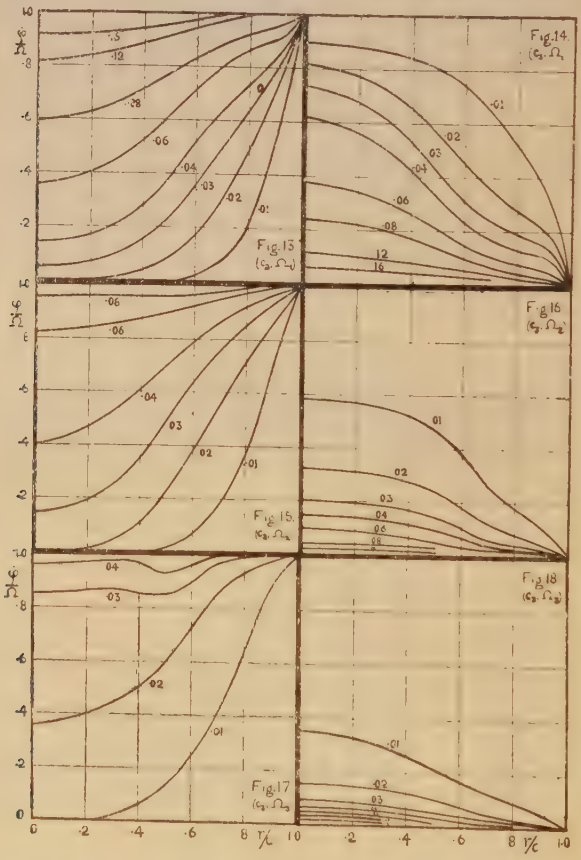
It is particularly noticeable that as the constant rotation Ω of the cylinder is increased, the departure from theory becomes more marked.



Only three figures were obtained for the small cylinder, because the agreement with the theoretical curves in the three cases omitted will be practically exact.

The curves for the "stopping" experiments show a much greater departure from theory than the curves for the "starting" experiments. This is due to the break-up of the regular motion, owing to instability at the fixed outer wall. Except with the high speed and the large cylinder,

the motion on starting, on the other hand, appears to the eye to be without appreciable irregularity, and it is very striking to see the sharp dividing-line between quiescent liquid in the centre and rotating liquid on the outside. This dividing-line, represented by the steep part of the curve in the figures, slowly moves towards the centre but becomes indistinct some distance from it. It is best seen at the higher speeds when



the velocity-gradient is greater. Its rate of travel depends on the value of Ω.

In the case of the large cylinder starting at the high speed, large secondary eddies 3·5 cm. across were often observed just inside the cylinder wall a few seconds after starting the motor. These soon died out when the velocity-gradient became less, and thereafter the motion travelled in towards

the centre regularly. The effect of these eddies was observed to cause a shifting of the curve for $\dot{\phi}/\Omega$ in the direction of a greater $\dot{\phi}$ for the same value of $\nu t/c^2$. In the same case, secondary eddies were observed about the ·4 circle when the motion had reached the centre. Although mean velocities were recorded (median line of the band of points), the effect is shown in fig. 17 by the wavy appearance of the two upper curves.

When the cylinder is stopped, the water continues to rotate until the irregular motions, generated near the cylinder wall, have had time to extend inwards. Small eddies then travel about, and the central axis of rotation wanders considerably and often seems to disappear temporarily amid cross-currents. The motion is very irregular except at the low speed, and even in this case some irregularity always remains. The lycopodium particles do not follow the circles for very long, and are usually moving at an angle to them.

With the large cylinder at the high speed, the velocity immediately after stopping the cylinder seemed to give stability and to aid in preserving the circular character of the motion: but when the kinetic energy had somewhat diminished, eddying became more noticeable.

On the curves mentioned in § 3 in which $\dot{\phi}/\Omega$ is plotted against $\nu t/c^2$, the bands of points are much narrower in the "starting" experiments than in the others, and determine the position of the median line easily to ·001 in the value of $\nu t/c^2$ in most cases. For the "stopping" curves, the limit of error may be two or three times this occasionally. One noticeable effect is that the band is narrow when the velocity-gradient has a considerable value, *i. e.* when the curves in the figures slope steeply. In these cases considerable momentum is being transferred through the water, and there will be considerable shearing stress and vorticity, and the stability might therefore also be considerable. As soon as the velocity-gradient becomes small, the band of points broadens. For example, in the "starting" curves the bands are sometimes very narrow until the value of $\dot{\phi}/\Omega$ has risen to 0·9, when they broaden out. Conditions seem to favour irregularity at the centre (axis) of the cylinder where the velocity-gradient vanishes. On the axis the stability is a minimum. In the "stopping" curves the bands are narrower the greater the angular velocity Ω, *i. e.* the greater the vorticity of the water, especially near the cylinder wall where the instability originates.

The observational curves show that viscosity alone is not

sufficient to account for the effects, except for small values of c and Ω, *i. e.* for long, narrow cylinders and slow speeds. A little dye introduced into the rotating water shows no signs of any minute eddying or micro-turbulence ; and so we must look for currents in the water as the cause of the discrepancy, which is obviously the case in the "stopping" experiments. The formation of large eddies in large bodies of fluid seems to be due chiefly to the interaction of two local currents, or to low pressure caused by an obstacle or a sink, and not to the slower processes of viscosity.

If we attribute the deviation from theory to an ignored increase in the kinetic viscosity ν, we find that when the large cylinder is stopped at the high speed, the increase would have to be represented by a factor exceeding 10 in value nearly everywhere, while the value would lie between 50 and 100 on $r/c=0{\cdot}3$ shortly after stopping. With the middle cylinder, stopped at the high speed, the factor has about half these values ; and with the small cylinder, stopped at the same speed, the factor ranges from 1·3 to 3·0. In the starting experiments the factors are nearly unity, but they are meaningless here as the motion is not turbulent.

Some earlier experiments illustrate the instability of the stopping experiments. In these an inner cylinder was rotated coaxially with a fixed outer one. As is well known, it was found that at no speed of rotation of the inner cylinder was it possible to set the water moving in circular paths, owing to the eddies which were continually thrown off. The slower the speed of rotation the more conspicuous were the eddies, especially on the borders of the outer, more slowly-moving water. Measurements of the angular velocity showed a large departure from theory, the inner parts rotating more slowly and the outer parts more rapidly than the theory indicates. The effect of the travelling eddies is thus to make the angular velocity more like that of a rigid body. When the speed was very great (2500 r.p.m.) the kinetic energy seemed to give stability to the water. A whirlpool formed next the inner cylinder, and a large oscillation was presently set up in the form of a wave with its crest along a radius of the outer cylinder and its trough on the other half of the same diameter.

Some thick, very viscous oil residues, when rotated in a cylindrical tin about 15 cm. in diameter, acquired the full velocity on starting (36 r.p.m.) in something less than 4 seconds, and came to rest in the same time when the cylinder was stopped. Only a slight displacement of the oil occurred, the surface being momentarily roughened with fine lines like cracks.

II. *The Elements of Geometry.*

By Norman Campbell, *Sc.D.**

Summary.

It is maintained that the geometry of Euclid is best interpreted as an attempt to deduce as many important propositions as possible from the assumption that length, angle, area (and perhaps volume) are magnitudes universally measurable by the methods that are actually employed in experimental physics. All his chief propositions (in so far as they are true) can be deduced from that assumption without any other.

This view is supported, not by a detailed analysis of the Elements, but by a very summary sketch of the laws that must be true if the assumption is to be acceptable. In a sequel it is hoped to discuss similarly the foundations of another branch of experimental geometry with which Euclid is not directly concerned—namely the geometry of position, which involves the concept of "space."

---

1. There was formerly much discussion whether geometry was an experimental or a mathematical science. It is now generally agreed that there are two closely connected sciences, one mathematical and one experimental. The former, which has been defined as the study of multi-dimensional series, consists of a logical development of ideas which have no necessary dependence on the experience of the senses. It does not consist of laws and cannot be proved or disproved by experiment; it can enter into relation with experimental science only through theories and by suggesting hypotheses which, interpreted suitably, predict laws. The formulation of such theories, in which Minkowski was the pioneer, is one of the most striking features of modern mathematical physics. The experimental science, on the other hand, is meaningless apart from experience, and its propositions are true or false according as they agree or disagree with experiment. They are the very fundamental laws which involve only the geometrical magnitudes such as length, angle, or area. It may be noted in passing that the laws predicted by geometrical theories are not in general geometrical laws, but involve electrical, optical, or dynamical concepts.

* Communicated by the Author.

The mathematicians who have recently taken over from the philosophers the task of teaching experimenters their business have decided that only the mathematical science is properly termed geometry. In support of their claim they appeal to the authority of the Greeks, and thereby imply that Greek geometry is mathematical and not experimental. This implication raises questions of scientific interpretation and not of mere convenience in nomenclature. For the matter cannot be decided by inquiring what Euclid (for example) thought he was writing about: it is admitted that, as an exponent of mathematical geometry, he was guilty of errors; and, if he was capable of error, he may have been wrong as to the nature of his assumptions and of his arguments. If we are justified today in confining the term to one study rather than another, because that term was used by Euclid, it can only be on the ground that Euclid's propositions and his methods of proving them are closely similar to those employed today in that study.

If this test is applied, geometry is an experimental science. For whereas the Elements is utterly different from anything modern mathematical geometers produce, it is, judged by modern standards, quite a creditable attempt at an exposition of experimental geometry. It can be regarded broadly as an attempt to deduce as many important laws as possible from the single assumption that length, area, angle, and (less definitely) volume are magnitudes, universally measurable by the methods which are actually employed in experimental physics, or to which the methods that are actually employed would be referred if doubt arose concerning their validity. *Nothing is assumed but that every straight line has a length, every pair of straight lines an angle, and every plane surface an area.* The definitions, axioms, and postulates should then be statements of the laws by virtue of which measurement is possible. It is admitted that the attempt is not wholly successful; but its faults, or many of them, are readily explicable: the author has not to be represented (as he must be if he is an exponent of the mathematical science) as constantly straining at gnats and swallowing camels.

Such a view can be established only by a detailed and tedious criticism which, in so far as it concerns Euclid's intelligence, is not of scientific interest. In place of it will be offered a very summary sketch of the fundamental notions and laws of experimental geometry and sufficient comparison of them with Euclid's assumptions to suggest that on them might be founded a deduction, by methods

very similar to those that he employs, of the propositions which he actually states. References are throughout to Todhunter's edition.

(2) But two preliminary questions must be asked. First, can an experimental science be deductive at all? Certainly it can. A deduction from a law is an application of that law in particular circumstances which were not examined when it was formulated. If, after examining the sides of squares and of triangles, I assert the general law that all straight lines have measurable lengths, and then, without further experiment, assert that the diagonals of squares, which are also straight lines, are also measurable, I am making a deduction. It may be true that there is something precarious about the results of such deduction—that question is not raised here,—but the deduction itself is quite unexceptionable; the falsity of the conclusion is definitely inconsistent with the truth of the premises. If doubt is raised concerning the conclusion, the ultimate means of resolving it is by experiment; but experimental science, in the hands of its greatest exponents, consists in asserting such general laws that doubt does not arise concerning the results of deduction based on them.

The second question is whether there are truly laws which make measurement possible. The question is discussed at length in my 'Physics,' Part II., the results and nomenclature of which will be used freely in what follows. But there is one matter which may receive special mention here, because it is concerned with "incommensurables," which are often (but falsely) believed to be of especial importance in geometry. Measurement is possible when, by means of definitions of equality and addition, a standard series of the property in question can be established, starting from some arbitrary unit, such that any system having the property is equal in respect of it to some one member of the standard series. Now (it might be argued) such measurement is not possible for length, because the diagonal of a square cannot be equal to any member of a standard series based on the side as unit; indeed that result is actually proved by Euclid. Consequently it is patently absurd to pretend that Euclid's propositions can be derived from an assumption, namely that measurement is possible, which is inconsistent with its conclusions.

One method of escape from this difficulty may be mentioned, although it will not be adopted. A slight

amendment in the thesis might be made, and it might be said that Euclid's assumption is that the laws are true which would make measurement possible if there were no incommensurable lengths—for these laws, though necessary to measurement, may not be sufficient. But the difficulty vanishes entirely, if it is remembered what is meant by "equality" in experimental measurement. When it is said that A is equal to B, it is meant that there is no possible means of deciding which of the two is the greater. If then I say that the diagonal of a square is $\sqrt{2}$ times the side, I mean that, if I measure the diagonal in terms of the side as unit, there is no means of deciding whether the value obtained, when multiplied by itself according to the multiplication table, will be greater or less than 2. That statement is not in the least inconsistent with my assigning to particular diagonals values of which the square is not 2; it is only inconsistent if a law can be found by which I can tell in particular cases whether the square will be greater or less than 2. My assertion is that there is no such law; and that assertion is true. In its application to all magnitudes except number, equality must be interpreted in this, slightly statistical, sense.

3. There is then no preliminary objection to the view that Euclid's propositions are deductions from the laws in virtue of which the geometrical magnitudes are measurable. We now proceed to ask what those laws are.

Geometrical conceptions are derived ultimately from our immediate sensations of muscular movement, just as dynamical conceptions are derived from our sensations of muscular exertion and thermal conception from our sense of hot and cold. We have an instinctive and indescribable appreciation of differences in *direction* of various movements; we appreciate that one direction may be *between* two others; and if other sensations (*e. g.* those of hot and cold or rough and smooth) vary with movement along a certain direction, we appreciate that of the varying sensations some are *between* others. The notions of direction and of the two kinds of betweenness are the foundations of geometry. It is a vitally important fact that there is an intimate relation connecting betweenness determined by one kind of muscular motion (*e. g.* that of the hand) and that determined by another (*e. g.* that of the eye). The relation is much too complex for any account of it to be attempted here; but it is only because it exists that "space" explored visually or by our different limbs is always the same.

The fundamental notions give rise to those of *surfaces* and *lines*. Surfaces are connected with the fact that a sensation may be unaltered by movement in any of a certain group of directions (which are said to be in a surface and to characterize it), while it may be altered by any movement in any direction not in this group (directions away from the surface). Of lines there are two kinds, which will be termed respectively "edges" and "scratches." Edges arise from the fact that the group of directions characteristic of a surface may change suddenly at some part of it. It is a matter of convenience whether the parts characterized by different directions are spoken of as different surfaces or as parts of the same surface: we shall adopt the second alternative. Scratches arise from the fact that, while the directions characterizing a surface are unaltered, the sensation the occurrence of which distinguishes "in the surface" from "out of the surface" may change suddenly. Some, but not all lines, are such that the whole of them lies along a single direction. *Points* are of little importance in the earlier stages of geometry; they arise from the fact that two lines may have a part in common. Two points, both on the same line, are termed the ends of the part of the line between those points.

The recognition of surfaces and lines is the first step towards geometry. Euclid attempts to give an account of them in Defs. 1, 2, 3, 5 of Book I., which are the least successful part of his treatise. The account given of them here is no better than Euclid's for the purpose of conveying a notion of them to one who does not possess it already; but since there are no such persons, the objection is not serious. But our account is better in drawing attention to the notions that are fundamental in geometry and in not assuming familiarity with conceptions, such as length, which are necessarily subsequent.

4. Some surfaces, but not all, when subjected to muscular force undergo only such changes as can be compensated by a suitable movement of the whole body; if such a movement is made, the group of directions characterizing the surfaces is restored. In other words, such surfaces can move without alteration of form; they provide the original and crude conception of a *rigid body*. By means of the motions of rigid bodies, it is sometimes possible to bring parts of two previously distinct surfaces into *contiguity*, so that there is nothing *between* those parts. In particular, edges, or parts of edges, can often be brought into such contiguity. Scratches

can be brought into contiguity with edges, and, in a sense, into contiguity with other scratches; but the criterion of contiguity in the last case is much less direct and requires methods involving something other than the simple perception of nothing between.

The recognition of the possibility of contiguity is the second step towards geometry and leads immediately to the third, which consists in the establishment of a definite criterion for a *straight line*. A crude criterion is provided by direct perception: a young child knows the difference between a straight and a bent line by simply looking at them; the recognition seems to depend on the fact that a straight line is all in one direction and is symmetrical with regard to the unsymmetrical directions of left and right or back and front. The crude criterion is stated as well as it can be in Euclid's Def. 4. But contiguity provides a much more stringent criterion, which in the first instance is applicable only to edges and not to scratches. Two edges are straight if, when two portions of one are brought into contiguity with two portions of the other, all the portions between these two portions are also in contiguity, however the contiguity of the first pairs of portions is effected. It appears as an experimental fact, that if A, B and C, D are two pairs of straight edges according to this criterion, C is also straight if tested against A; accordingly an edge can be called straight independently of the other member of the pair on which the test is carried out. A scratch is straight if it can be brought into complete contiguity with a straight edge. These facts are stated in Axiom 10.

Other definitions of a straight line are sometimes offered: *e. g.*, (1) an axis of rotation, (2) the shortest distance between two points, (3) the path of a ray of light. (1) is almost equivalent to that stated here; (2) will be noticed presently; (3) is not accurately true (*i. e.*, if it is adopted, the familiar propositions about straight lines are not true), but it is important as an approximation for comparatively rough measurements.

A *plane* surface (or, according to our usage, part of a surface) is then defined as in Def. 7. It can also be defined by the complete contiguity of three pairs of surfaces; but the contiguity of surfaces is not easy to describe accurately. Such a definition is, however, actually used in making optical flats and surface plates; if it were adopted, it would still be necessary to introduce the fact that it agrees with our definition, in order to measure angle. The

conception of the contiguity of surfaces is not actually required, except perhaps for the measurement of volume. (*Cf.* § 11.)

5. The third step places us in a position to introduce measurement and the three fundamental magnitudes, *length*, *angle*, and *area*. For fundamental measurement we need definitions of equality and addition, such that the law of equality and the two laws of addition are true. The choice of unit may be left out of account; for, with geometric magnitudes, the laws are true whatever unit is selected. The law of equality is Axiom 1; the first law of addition is Axiom 9. Axioms 2–7 are together very nearly equivalent to the second law of addition (which may be stated roughly in the form that the magnitude of a sum depends only on the magnitudes of the parts). Axiom 8 is an attempt to compress the definitions of equality for all three magnitudes into a single sentence; it is better to separate them. Euclid fails to give any definition of addition; he does not tell us how the "whole" is to be related to the "parts" in order that it should be greater.

6. We will now take the magnitudes in turn. For the *length* of a straight line the necessary definitions are:— (1) Two straight lines are equal in length if they can be placed so that when one end of the first is contiguous with one end of the second, the other ends are also contiguous. (2) The length of the straight line AB is equal to the sum of the lengths of the straight lines CD, EF, if they can be placed so that C is contiguous with A, F with B, D with E and with some part of AB between A and B.

These definitions, like all similar definitions of magnitudes, are satisfactory and are subject to the necessary laws of equality and addition only if certain conditions are fulfilled. The conditions are described by saying that the surfaces in which the straight lines lie must be those of rigid bodies. This is a definition of a rigid body: a rigid body is something which (like a perfect balance) is determined by the satisfaction of the conditions for measurement*. Rigid bodies according to this test include many of those which satisfy the crude test of § 4, though they include others (*e.g.*, surveyors' tapes used as surveyors use them) which do not satisfy that test. In virtue of the fact that rigid bodies are necessary to measurement, the

* *Cf.* H. Dingler, *Phys. Zeit.* xxi. p. 487 (1920).

branch of geometry with which we (and, according to our view, Euclid) are concerned may be fitly described as the study of the surfaces of rigid bodies. It is thus distinguished from a wholly different branch of geometry, with which we are not here concerned, that is not confined to rigid bodies; this is the geometry of position.

It is important to notice that not all pairs of straight lines can be brought into contiguity, and that the law of equality cannot therefore be tested universally. It might have turned out that there was some material difference between those which can and those which cannot be brought into contiguity with a given line; and that if we assumed that the law of equality is universally true, we should be led to inconsistencies. It is an experimental fact that no such inconsistencies do arise when we extend our definition of equality so that lengths are equal when they are equal to to the same length, although they cannot be brought into contiguity with each other. This is, of course, one of the most important laws that make measurement possible. A similar remark applies to all the geometric magnitudes and need not be repeated.

7. The length of lines that are not straight can be measured *approximately* as fundamental magnitudes by means of flexible but inextensible strings. But the laws of such measurement are not strictly true, because (as we say now) no string is infinitely thin and the surface never coincides with the neutral axis. Another possible way, perhaps more accurate but of limited application, would be to roll curved edges on some standard edge, which need not be straight. But in truth there is no perfectly satisfactory way of measuring fundamentally the length of curved lines. All the measurements which we make on them are derived from measurement of straight lines; they involve numerical laws between fundamentally measured magnitudes. One of these laws is that the perimeters of the circumscribed and inscribed regular polygons tend to a common limit as the number of sides is increased. That law is therefore a law of measurement if curved lines are to be measured.

The question whether curved lines can be measured fundamentally is important, because, if they could be, it would be possible to define a straight line as the shortest distance between two points. (The definition would have to be put in some other form, since distance, a conception belonging to the geometry of position, implies the measurement of length.) But since they cannot be, that

definition must be rejected : it must be regarded merely as a generalized form of Prop. I. 20.

8. *Angle* is the measure of the crude conception of direction. The following are the definitions of equality and addition for the angle between two intersecting straight lines :—The angle between two straight lines A, B is equal to that between C, D if it is possible to bring A into continguity with C and B with D. The angle between A, B is the sum of the angles between C, D and E, F, if, when A is brought into contiguity with C and D with E, D lying between C and F and in the same plane with them, F can be brought into contiguity with B. These definitions are satisfactory only if the straight lines are in rigid bodies; or, in other words, there are surfaces which satisfy the conditions for the measurement of length and also those for the measurement of angle.

But even if the surfaces are those of rigid bodies, the definitions are not wholly satisfactory and the laws of measurement not entirely true. We must distinguish angles according as the two straight lines which they relate are or are not prolonged on both sides of the common point : the latter class may be termed "corners," the former "crossings." Angles between edges are always corners ; those between scratches may be either corners or crossings. If we try to include both corners and crossings in the same class as a single magnitude, the law of equality is not true ; for two corners which are both, according to the definition, equal to a crossing may not be equal to each other ; as we say now, one angle may be the supplement of the other. But if we treat corners and crossings as separate magnitudes this difficulty disappears ; the law of equality is true for either taken apart from the other. Actually we take corners only as magnitudes ; crossings we measure by the corners with which they can be made contiguous. Each crossing then has four angles (*i. e.* corners) associated with it. It is an important experimental fact that the "opposite" angles are equal ; it is best taken as a primary law, instead of being proved from other axioms as in Prop. I. 15. It is a law of measurement, because if it were not known, we should need four and not two angles to measure a crossing ; it is thus inherent in our system of measurement.

But though the law of equality is now true, the first law of addition is false ; it is false for both corners and crossings. The whole which is the sum of the parts may be equal to

one of the parts: *e. g.*, if both of two parts, being corners, are what we now call 120°. Some kind of spiral space can be imagined in which the law would be true; but actually it is very important that it is false. For, apparently inseparable from its falsity is the fact that the angle between two portions of the same straight line can be measured and given a finite value in terms of a unit which is the angle between two intersecting lines. This fact is described by the assertion that there are right angles and that a perpendicular can be drawn to any straight line from any point in it, a right angle being defined as in Def. 1. 10. (Axiom 1. 11 follows from this definition, regarded as an existence theorem, and our axiom Prop. I. 15.) Since the existence of right angles is vital to geometry, we cannot avoid the falsity of the first law of equality by some alteration of the definition. We can only recognize that the law is true in some conditions, and be careful to apply it in deduction only when it is true. It is true when all the lines making the added angles lie on the same side of (or contiguous with) a single straight line passing through their common point; this condition can be expressed, though with some complexity, in terms of the fundamental notion of between. Thus, in proving Prop. I. 16 we need to know that CF and CD both lie on the same side of AG. This law, and perhaps others of the same nature, are laws of measurement, defining the conditions in which angle can be measured uniquely. They require explicit mention.

The ambiguity which the falsity of the first law of addition introduces into numerical measurement is removed by certain conventions. These need not be considered here for we are not assigning numerical values.

If the length of curved lines were measurable fundamentally, angle might be measured as a pure derived magnitude, *e. g.* by the ratio of the arc to the radius of a circle in virtue of the numerical law, established experimentally, that the arc is proportional to the radius. But since curved lines cannot be so measured, we must take angle to be fundamental. We cannot use right-angled triangles with straight sides to measure angle as derived, because we need fundamental measurement to determine what angles are right. Of course we might define for this purpose a right angle as an angle between some two lines arbitrarily chosen as standard; but such measurement would be intolerably artificial and nothing whatever could be deduced from such a definition.

9. Euclid's definition of *parallel lines* must be rejected entirely, for, since all plane surfaces are limited, the criterion suggested is inapplicable. Since the crude definition of parallelism is similarity of direction, we may try to define parallel lines as those which being in the same plane make the same angle with any third line. We thereby imply the axiom of parallels in the form (Prop. I. 29) that such lines which make the same angle with one straight line make the same angle with any other; we imply also that the angles which are to be equal are the "exterior" and "interior" opposites or the "alternate" angles, since if the interior angles are compared the proposition is not true. But the definition is not very satisfactory; for, when the lines are edges, there is not always an exterior or an alternate angle. It is better to adopt the substance of Axiom I. 12 as a definition, and to say that lines in one plane are parallel when the sum of the interior angles is equal to two right angles. This much abused axiom seems to me a very ingenious way out of a real difficulty. We then assert the axiom of parallels in the form (implied by I. 32) that if any two straight lines in a plane are cut by any third line, the sum of the interior angles is the same for all third lines. The merit of this axiom is that it indicates clearly that the "axiom of parallels" is really something concerning all straight lines in a plane and not only parallel lines, and that parallel lines are merely a particular case of other pairs of lines. The propositions that parallel lines never do intersect and that the angle between them is zero follow immediately.

The axiom of parallels is a law of measurement because it is involved in the measurement of the angle between lines which do not intersect. Its use for this purpose requires that at some point of a straight line it should always be possible to place a straight line parallel to a given straight line. This proposition is not true for concave surfaces, but the complexities arising from this failure and the means of avoiding them may be left for the present; they are dealt with more naturally in connexion with "space." If the axiom were not used, we could not by our present methods measure the angle between non-intersecting straight lines: first, because the definition of equality given above, though sufficient for such lines, is not necessary: second, because the definition of addition is wholly unsatisfactory.

There has been so much discussion of the necessity of the

axiom of parallels that the matter requires rather more consideration. Two questions are involved. First, would it be possible to measure the angle between non-intersecting lines without assuming some proposition logically deducible from the axiom? It would be if, and only if, some property, common to all lines between which the angle is the same, can be found which is determinable by direct experiment not involving parallel lines. There may be such a property, but I have not been able to think of it. Second, if the axiom were not actually true—but we may stop there. In a pure experimental science, there is no sense in asking what would happen if the world were other than it actually is. Theory is necessary to give such a question a meaning, by suggesting what might remain unaltered during the change. For our present purpose the axiom is as necessary as any other of those we are considering.

10. *Area* is distinguished from all other fundamental magnitudes because the definitions of equality and addition are inseparable. They may be expressed thus. The areas of two bounded plane surfaces are equal if (but not only if) their boundaries can be brought into complete contiguity with each other or with the same third boundary. (A bounded surface is a part of a surface which includes all portions which can be traversed without crossing the boundary line.) The area of A is the sum of the areas of B and C, if when parts of the boundaries of B and C are brought into contiguity with each other, the remaining parts of the boundary can be brought into contiguity with the boundary of A. In virtue of the fact that parts of the boundaries of two surfaces can be brought into contiguity in many different ways, there may be many different bounded surfaces, of which the boundaries cannot be made contiguous, which are the sum of the same bounded surfaces. If the measurement of area is to be satisfactory, these surfaces must also be deemed to have equal area, and the definition of equality must be extended correspondingly. With this extension the laws of equality and addition are true, and the measurement is satisfactory.

In order that all bounded plane surfaces should have areas, some rule must be found for choosing the shape of the members of the standard series and for grouping them in such a way that some sum of them is equal to any area. We use for this purpose rules based on the axiom of parallels, and that axiom is therefore again a law of the

measurement of area. The rule might possibly be dispensed with, if we were prepared to spend unlimited time in selecting by trial and error shapes for the members of the standard series which fulfil the necessary conditions; but actually we could never measure area except by making use of similar figures, the production and properties of which depend wholly on the axiom of parallels. Further, it is the use of that axiom which enables us nowadays to calculate area from the linear dimensions of a surface without resorting at all to fundamental measurement. But of course all the numerical laws on which that calculation depends have to be established by means of fundamental measurement. It is only by defining area as we have done, and assuming the axiom of parallels, that we can prove by deduction that the area of a rectangle is proportional to the product of its sides, or equal if the units are suitably chosen.

The areas of surfaces that are not plane cannot be measured fundamentally, even to the extent that the length of curved lines can be. For there are no inextensible surfaces which can be brought into contiguity with surfaces of any curvature. Measurement of curved area is always derived and estimated by the limit of the circumscribed polyhedra as the number of their sides is increased. But the whole matter is obscure, because it is much more difficult to establish experimentally that there is a limit or to say what the limit is; for there is here no inscribed polyhedron tending to the same limit. There is singularly little experimental evidence for the assertion that the area of a sphere is $4\pi r^2$, and there is great difficulty in saying exactly what we mean by such an assertion; curved area is almost always a hypothetical idea and not an experimental magnitude at all.

11. *Volume* is a property of complete surfaces. Since complete surfaces can never be brought into complete contiguity, volume cannot be measured fundamentally by any process at all similar to those applicable to the magnitudes we have considered so far. Volume is measured (1) as a fundamental magnitude by means of incompressible fluids, or (2) as a derived magnitude by means of the lengths and angles characteristic of the surface. The second method depends upon numerical laws established by means of the first. In certain cases these laws can be related closely to other geometric laws by means of the following propositions:—(1) Two complete surfaces with equal dimensions,

i. e. with equal lengths and equal angles between them, have equal volumes. (2) If two complete surfaces have each one part plane, and the boundary of the plane part of one can be brought into complete contiguity with the plane part of the other, then the complete surface which has dimensions equal to that of the complete surface so formed has a volume equal to the sum of the volumes of the original surfaces. These propositions could be used as definitions of equality and addition in a system of measurement, which would be independent of the measurement of length and angle (and therefore not derived), because it involves only equality, and not addition, of length and angle. But it is of limited scope and, in particular, would not permit the measurement of the volumes of curved surfaces. Since we do undoubtedly attribute a meaning to the volume of such surfaces, in a way that we do not to their area, measurement by incompressible fluids, which is not geometric, cannot be wholly avoided. But the propositions, which are those on which Euclid bases his treatment of volume, are actually used in modern practice, and are therefore regarded permissibly as laws of measurement.

12. In deducing Euclid's propositions from the laws of measurement of these magnitudes, subsidiary laws are required, corresponding roughly to his postulates, expressed and implied. First, we need "existence theorems" corresponding to each of the definitions; for example, the definition of a plane surface justifies the conclusion that a straight edge can be placed contiguously to any two portions of such a part of a surface. Second, we need the assumption that we can make an object having a magnitude equal to that of any object presented to our notice. All these propositions are laws of measurement: the first group, because all definitions in experimental science are nothing but existence theorems; the second, because it is implied in the fact that we can make a standard series by which we can measure any magnitude.

Euclid's three expressed postulates are all untrue. I cannot "draw a straight line" from this room to the next when the door is closed. Moreover his constructional propositions, closely connected with the postulates, are unsatisfactory because they are all directed to the drawing of scratches, rather than to the making of edges. The hypothetical experiments by means of which the deductions are effected are carried out much more easily with edges

than with scratches : and if any of the propositions were doubted and put to the test of experiment, it would certainly be by means of edges : the extension to scratches would be by means of the contiguity of edges with them. Euclid's methods here undoubtedly indicate that he is leaving, perhaps consciously, the realities of experimental science for the pure ideas of mathematics. But he has made so little progress towards the new peak that, if he is to be restored to safety, it is far easier to drag him back to that which he has never left completely than to guide him through the bog in which the two sciences are confused to the very distant goal.

13. Only a few disconnected remarks will be offered here on the process of deducing the Euclidean propositions from the fundamental laws that have been sketched. Of course, we should employ the "application" (or contiguity) method of Prop. I. 4 wherever possible, instead of trying to avoid it : for it is based directly on the fundamental notions. Again, we should not commit Euclid's error of supposing that strictly similar triangles can be brought into contiguity ; we should apply the mirror image first to one triangle and then to the other. There would be no need to introduce area to prove Prop. I. 47. A Greek writer was forced to do so, because, not being familiar with the multiplication table, he could describe in no other way the relation between a number and its product by itself. We should proceed from Prop. I. 34 to Book VI. and prove Prop. I. 47 by drawing the perpendicular from the right angle to the hypoteneuse and using the relations of similar triangles, treated by algebra. For nowadays, since we admit no incommensurable magnitudes, we can dispense altogether with Euclid's very beautiful and ingenious subtleties about ratios. A ratio in experimental science is nothing but a value taken from the multiplication table, which is established by the measurement of number, *i. e.* by counting. The laws of the measurement of number are involved in those of the measurement of every "continuous" magnitude.

April 22, 1922.

III. *On the Rotation of Slightly Elastic Bodies.* *By* Dorothy Wrinch, *D.Sc., Fellow of Girton College, Cambridge, and Member of Research Staff, University College, London* *.

THE change in dimensions of a slightly elastic body due to rotation is a question of some practical importance, and does not appear to have received any systematic treatment. In the theory of elasticity, the displacements of a point of the body are of course discussed and the displacements of the points of the boundary determine the increase of dimensions. But the problems of elasticity which are of interest mainly from the point of view of increase of dimensions, rather than of the distribution of stress in the material, can rarely be solved by the current methods or appear only as special cases of a general mode of analysis. Even the simple problem of a circular cylinder of finite length, rotating about its axis, has not yet admitted an exact solution, though an approximate solution, which becomes valid when the cylinder is of infinite length, has been given by Chree. When the cylinder has a finite length, the surface condition of zero traction over the curved surface is violated, and instead of this traction becoming zero at all points on the surface, only its average value over the surface is zero. The results for the case of an infinite cylindrical annulus do not appear to be on record, and they are interesting on account of their marked divergence from those which belong to the complete disk.

In the present paper we group together some of the simpler and more interesting solutions of problems of this type, including those of the infinite circular cylinder and the infinite cylindrical annulus. These specific problems are solved to any degree of approximation and for a non-uniform distribution of density. The analysis is simpler than is usual, for it does not seem necessary to treat these comparatively simple problems as special cases of general theory, and it is desirable, at least in the interests of the engineer or physicist, that a fundamentally simpler treatment should be placed on record. It also seems possible that such solutions may be of interest with regard to scientific instruments of great precision, in which some portion of the apparatus is in rotation, or, on the larger scale, in problems of practical engineering. Although no

* Communicated by the Author.

novelty attaches to some of the earlier results, it seems desirable to include them.

The simplest problem of this nature is, of course, that of the thin circular hoop rotating about its centre. When such a hoop of radius a and density ρ is spun round its centre with constant angular velocity ω the value of T, the tension per unit length in the hoop, is well known. For an element ds of its length has an acceleration $a\omega^2$ inwards, and the resultant of the tensions at its ends is Tds/a per unit area inwards. Hence the equation of motion is

$$Tds/a=\rho a\omega^2,$$

giving

$$T=\rho a^2\omega^2.$$

If, however, the hoop is slightly elastic, and λ the value of Young's modulus for the material of which the hoop is made, and r the radius of the hoop when in motion, the equation of motion of the stretched element ds becomes

$$T/r=r\omega^2 \,.\, \rho a/r.$$

Applying Hooke's law to the stretched element, we have,

$$T=\lambda(r-a)/a.$$

Hence eliminating T,

$$\rho r a\omega^2=\lambda(r-a)/a.$$

In practice λ is always large, and if we may neglect $1/\lambda$ and higher powers of $1/\lambda$ the appropriate value of r/a, which differs from unity by a quantity of order $1/\lambda$, is $1+\rho a^2\omega^2/\lambda$. The value of the tension to the same order is $\rho a^2\omega^2$.

The effect of a rotation is therefore to increase the radius a of the hoop to $a(1+\mu)$, where $\mu=\rho a^2\omega^2/\lambda$, a number depending on the density, the elasticity, and the radius of the hoop, and on the rate at which it is rotating.

As regards the practical order or magnitude of $\rho a^2\omega^2/\lambda$ the extension per unit length, we may take a steel wire for which λ is about $2{\cdot}12\times10^{12}$ dynes per square centimetre, and ρ is about 7·5. In order that Hooke's law may hold, the extension per unit length must not exceed 10^{-3}, roughly speaking. If the velocity of a point on the rim is in the neighbourhood of $1{\cdot}9\times10^4$ cm. per second—which is approximately the case in a twenty-foot flywheel making two hundred and fifty revolutions a minute—we find that the extension per unit length is about $7{\cdot}9\times10^{-4}$, which comes within the limits of applicability of Hooke's law, and that

the actual increase in the radius is about a fifth of an inch. In this case the tension is about $1{\cdot}6 \times 10^9$.

It is further evident that $\sqrt{n\lambda/\rho}$ is the largest velocity if an extension of more than n per unit length is to be avoided. When the elastic limit for the material is known, this result can be used to give an upper limit to the velocity it is safe to use if risk of deformation of the hoop is to be avoided.

We may now proceed to the problem of a thin rod rotating about one end with uniform angular velocity.

Thin Rod Rotating about One End.

Let a_0 be the unstretched length of the rod, ω the angular velocity of rotation about one end O, ρ_0 the density when it is unstretched, and λ the value of Young's modulus for the material of which the rod is made. Let T be the tension in any section in the rod during the motion. Let the distances of the same particle at rest and in motion be x_0 and x. The density of the moving element dx is $\rho_0 dx_0/dx$ and its acceleration towards O is $x\omega^2$. The equation of motion of the element is therefore,

$$\delta T = -\rho_0 \delta x_0 \,.\, x\omega^2,$$

where, by applying Hooke's law to the element originally of length dx_0 and now of length dx, we have

$$T = \lambda(dx/dx_0 - 1).$$

Hence, eliminating T, we obtain the equation,

$$d^2x/dx_0^2 = -\rho_0\omega^2 x/\lambda.$$

The solution must give the value $x=a$ when $x_0=a_0$ if a is the length of the rod when in motion. Accordingly it is

$$x = a \sin(x_0 \sqrt{\rho_0\omega^2/\lambda}) \,/\, \sin(a_0 \sqrt{\rho_0\omega^2/\lambda}).$$

We may determine the value of a by means of the condition that the tension vanishes at the free end, which is given indifferently by $x=a$ or $x_0=a_0$. Thus,

$$a \sqrt{\rho_0\omega^2/\lambda} = \tan a_0 \sqrt{\rho_0\omega^2/\lambda}.$$

The equation relating the two corresponding positions of a

typical element when at rest and when in motion and the original length of the rod is therefore

$$x\sqrt{\rho_0\omega^2/\lambda}=\sin(x_0\sqrt{\rho_0\omega^2/\lambda})/\cos(a_0\sqrt{\rho_0\omega^2/\lambda}).$$

Neglecting the cube and higher powers of $1/\lambda$, we may replace this by the simpler form,

$$x=x_0+x_0\rho_0\omega^2(3a_0^2-x_0^2)/6\lambda$$

to the order $1/\lambda$. To the same order,

$$T=\rho_0\omega^2a_0^2\left[1-x_0^2/a_0^2\right].$$

The greatest extension is $\frac{1}{2}\rho_0\omega^2a_0^2/\lambda$, and this occurs at the end about which the bar is rotating. The tension is also greatest at this point and takes there the value $\rho_0\omega^2a_0^2$.

As an example of the actual magnitudes of the quantities in practical cases we may take a twenty feet steel bar, which, when rotating about one end two hundred and fifty or three hundred times a minute, increases in length about a tenth of an inch.

Rotation of an Infinite Elastic Circular Cylinder about its Axis.

Passing now to a simple problem in three dimensions, we take the case of an infinite elastic cylinder of circular section rotating about its axis. We may consider one of the circular sections of the cylinder and use polar coordinates. At any point (r, θ) let T_1 and T_2 be the transverse and radial tensions per unit length, and T_3 the axial tension. We shall consider the motion of the element of volume which when at rest is bounded by the surfaces $(z, z+\delta z)$, $(r, r+\delta r)$, $(\theta, \theta+\delta\theta)$. By the symmetry of the cylinder, the element when in motion will continue to be bounded by the surfaces $(\theta, \theta+\delta\theta)$: and since the cylinder is of infinite length, the element will continue to be bounded by the surfaces $(z, z+\delta z)$. Let ρ represent the radial dimension, so that $\rho-r$ is the radial extension at any point. Let a be radius of the cylinder and σ its density, when at rest; let ω be the angular velocity of the cylinder about its axis, and λ and μ the elastic constants for the material of which the cylinder is made.

The element of volume which we are considering is a parallelepiped of sides $d\rho$, $\rho d\theta$, and dz. The forces on our element of volume consist of (1) transverse tensions each

of magnitude $T_1 d\rho\, dz$—and these are equivalent, in the usual way, to a radial force towards the centre of magnitude

$$T_1 d\rho\, dz\, ds/\rho,$$

where $ds = \rho d\theta$; (2) of radial tensions

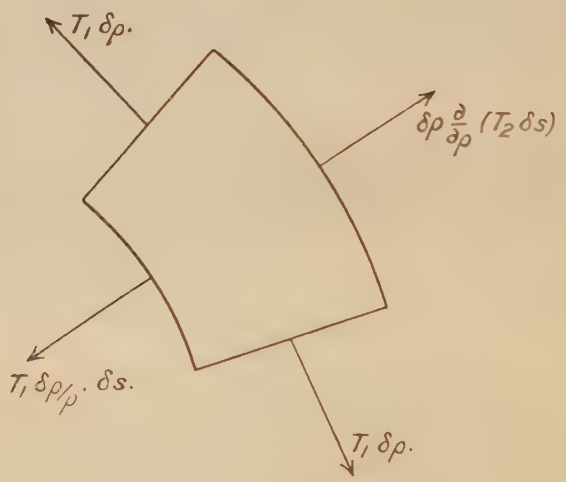
$$T_2 ds\, dz \quad \text{and} \quad T_2 ds\, dz + d\rho\, d(T_2 ds\, dz)/d\rho,$$

towards and away from the centre, which together give a force

$$d\rho\, d\theta\, dz\, d(\rho T_2)/d\rho$$

away from the centre; and (3) of longitudinal tensions each

Fig. 1.



of magnitude $T_3 d\rho\, ds$, in opposite directions. The resultant force then is simply

$$d\rho\, dz\, ds\, (d(\rho T_2)/dr - T_1)$$

away from the axis, and perpendicular to the axis of the cylinder. The acceleration of the element is $\rho\omega^2$ towards the centre and its mass is $\sigma r\, dr\, d\theta$, since we may, of course, treat the density as constant over the element of volume. We therefore have the equation of motion,

$$d(\rho T_2)/d\rho - T_1 = -\sigma r \rho \omega^2\, dr/d\rho.$$

By Young's law we can express T_1, T_2, and T_3 in terms of the extensions of the cylinder, in the well-known equations,

$$T_1=\lambda(d\rho/dr-1)+(\lambda+2\mu)(\rho/r-1),$$
$$T_2=(\lambda+2\mu)(d\rho/dr-1)+\lambda(\rho/r-1),$$
$$T_3=\lambda(d\rho/dr+\rho/r-2).$$

Putting these values of T_1 and T_2 in the above equation we obtain the result

$$\frac{d}{dr}\left[\rho(\lambda+2\mu)\left(\frac{d\rho}{dr}-1\right)+\rho\lambda\left(\frac{\rho}{r}-1\right)\right]$$
$$-\frac{d\rho}{dr}\lambda\left(\frac{d\rho}{dr}-1\right)-\frac{d\rho}{dr}(\lambda+2\mu)\left(\frac{\rho}{r}-1\right)=-\sigma r\rho\omega^2,$$

or

$$(\rho/r)\,d^2\rho/dr^2+(1/r)[(2\mu/r)(d\rho/dr)(\lambda+2\mu)+(\lambda\rho/r)(\lambda+2\mu)]$$
$$[d\rho/dr-\rho/r]=-\rho\sigma\omega^2/(\lambda+2\mu). \quad . \quad (1)$$

The value of μ or λ varies from about 8×10^8 grammes' weight per square centimetre for steel to about 4×10^8 grammes' weight for copper. The corresponding densities are about 7 and 9 respectively. Terms involving

$$\sigma/(\lambda+2\mu)$$

are therefore of a smaller order than those which involve coefficients of the form

$$2\mu/(\lambda+2\mu).$$

Let us write $\omega^2\sigma=q(\lambda+2\mu)$. Then putting

$$\rho-r=q\eta_1+q^2\eta_2+q^3\eta_3+\ldots.$$

we can obtain a value for ρ to any order of approximation which is required. Neglecting, first, all terms involving q^2, we have the equations for η,

$$\rho=r+q\eta_1\,;\quad d\rho/dr=1+q\,d\eta_1/dr\,;\quad d^2\rho/dr^2=q\,d^2\eta_1/dr^2\,;$$
$$q\,[1+q\,\eta_1/r]\,d^2\eta_1/dr^2+1/r\,[1+q\,(2\mu\,d\eta_1/dr)/(\lambda+2\mu)$$
$$+(\lambda\eta_1/r)/(\lambda+2\mu)+q\,(d\eta_1/dr-\eta_1/r)]=-qr(1+q\,\eta_1/r)\,;$$
$$q\,d^2\eta_1/dr^2\,(1+q\,\eta_1/r)+1/r\,[1+q\,(2\mu\,d\eta_1/dr)/(\lambda+2\mu)$$
$$+(\lambda\eta_1/r)/(\lambda+2\mu)]\,q\,(d\eta_1/dr-\eta_1/r)+qr\,[1+q\eta_1/r]=0\,;$$

and since we are neglecting terms in q^2, the equation for η_1 is

$$r^2\,d^2\eta_1/dr^2+r\,d\eta_1/dr-\eta_1=-r^3,$$

giving a solution of the form

$$\eta_1 = a_1 r + b_1/r - \tfrac{1}{8}r^3$$

and $b_1 = 0$, since η_1 is not infinite at the axis. The boundary conditions determine the constants a_1 and b_1; for the radial traction T_1 is given by

$$T_1 = (\lambda + 2\mu)\left(\frac{d\rho}{dr} - 1\right) + \lambda\left(\frac{\rho}{r} - 1\right)$$

$$= q\,[(\lambda + 2\mu) d\eta_1/dr + \lambda\,\eta_1/r],$$

and the radial traction must vanish over the two boundary surfaces. Thus

$$(\lambda + 2\mu)\,\overline{d\eta_1/dr}\,|_{r=a} + \lambda\,\overline{\eta_1/r}\,|_{r=a} = 0.$$

The constant a_1 is therefore determined by the equation,

$$(\lambda + \mu) a_1 = (2\lambda + 3\mu)\, a^2/8,$$

giving

$$T_1 = q\,[2(\lambda + \mu) a_1 - r^2/4(2\lambda + 3\mu)]$$

$$= q\,(2\lambda + 3\mu)(a^2 - r^2)/4r^2$$

and

$$\eta_1 = q\,(a_1 r - r^3/8)$$

$$= \frac{qr}{8}\,[(2\lambda + 3\mu)(a^2 + b^2)/(\lambda + \mu) - r^2].$$

The Effect of a Circular Hole in the Cylinder.

If the cylinder at rest has two boundary surfaces $r = a$, $r = b$ $(b < a)$ the solution stands in the form,

$$\eta_1 = a_1 r + b_1/r - \tfrac{1}{8}r^3,$$

and the conditions for zero radial traction on both the bounding surfaces yield

$$(\lambda + \mu) a_1 = \mu b_1/a^2 + (2\lambda + 3\mu) a^2/8,$$

$$(\lambda + \mu) a_1 = \mu b_1/b^2 + (2\lambda + 3\mu\ b^2/8.$$

These give

$$a_1 = (2\lambda + 3\mu)(a^2 + b^2)/8(\lambda + \mu),$$

$$b_1 = (2\lambda + 3\mu) a^2 b^2/8\mu.$$

Thus

$$T_1 = q\,[2(\lambda + \mu) a_1 - 2\mu b_1/r^2 - (2\lambda + 3\mu) r^2/4]$$

$$= q\,(2\lambda + 3\mu)(a^2 - r^2)(r^2 - b^2)/4r^2.$$

And the radial extension is

$$\rho - r = q\left[a_1 r + b_1 r - r^3 8\right]$$

$$= \frac{qr}{8}\left[(2\lambda+3\mu)(a^2+b^2),(\lambda+\mu) + (2\lambda+3\mu)a^2b^2/\mu r^2 - r^2\right].$$

The maximum radial traction occurs where

$$(a^2-r^2)(r^2-b^2)/r^2$$

is greatest, namely, at $r^2=ab$. The greatest radial traction therefore occurs on the cylinder whose radius is the geometric mean of the radii of the bounding surfaces of the cylinder. Its value here is

$$\rho\omega^2(2\lambda+3\mu)(a-b)^2/4(\lambda+2\mu).$$

The transverse tension is given by

$$T_2 = q\left[(\lambda+2\mu)\,\eta_1/r + 2\lambda\, d\eta_1/dr\right]$$

$$= q/4\left[(2\lambda+3\mu)(a^2+b^2) + (2\lambda+3\mu)\,a^2b^2/r^2 - (\lambda+2\mu)r^2\right]$$

$$= (2\lambda+3\mu)\frac{q}{4}\left[a^2+b^2-a^2b^2/r^2 - (\lambda+2\mu)r^2/(2\lambda+3\mu)\right].$$

The maximum value of this tension occurs on the cylinder with radius r_2 given by

$$r_2 = ab\sqrt{\{(2\lambda+3\mu)/(\lambda+2\mu)\}},$$

and there its value is

$$\frac{q(2\lambda+3\mu)}{4}\left[a^2+b^2-2ab\sqrt{(\lambda+2\mu)/(2\lambda+3\mu)}\right].$$

Finally, the tension parallel to the length of the cylinder is

$$T_3 = q\lambda(\eta_1/r + a\eta_1/dr)$$

$$= \frac{q\lambda}{4}\left[\frac{2\lambda+3\mu}{\lambda+\mu}(a^2+b^2) - 2r^2\right].$$

The longitudinal tension T_3 is therefore greatest on the inner boundary. Here its value is

$$\frac{q\lambda}{4}\left[\frac{(2\lambda+3\mu)a^2 + (\mu-2\lambda)b^2}{\lambda+\mu}\right].$$

In the well-known case of a solid cylinder, the longitudinal tension has its maximum on the axis, and its value there is

$$(2\lambda+3\mu)\,qa^2\lambda/4(\lambda+\mu).$$

The radial and the transverse tensions have also their maximum value on the axis, where they take the same value,

$$\rho\omega^2(2\lambda+3\mu)/4a^2(\lambda+2\mu).$$

The severest traction the cylinder is called upon to withstand is therefore

$$\rho\omega^2(2\lambda+3\mu)/4a^2(\lambda+2\mu),$$

and it occurs radially and transversely on the axis of the cylinder.

In the case of the cylindrical annulus, the facts are entirely different. The longitudinal tension reaches its maximum value

$$\rho\omega^2\lambda[(2\lambda+3\mu)a^2/(\lambda+\mu)-\mu b^2/(\lambda+\mu)]/4(\lambda+2\mu)$$

on the inner boundary. The transverse tension, in general, reaches its maximum value

$$\rho\omega^2(2\lambda+3\mu)\left[a^2+b^2-2ab\sqrt{(\lambda+2\mu)/(2\lambda+3\mu)}\right]/4(\lambda+2\mu),$$

on the cylinder $r=r_2$. And the radial tension reaches its maximum value

$$\rho\omega^2(2\lambda+3\mu)(a-b)^2/4(\lambda+2\mu),$$

on the cylinder $r=r_1$. As the ratio λ/μ varies, the place at which the maximum tension occurs varies. If the ratio b/a is sufficiently large, *i. e.* if

$$b\sqrt{\frac{2\lambda+3\mu}{\lambda+2\mu}}\geq a, \quad b/a\geq\sqrt{\frac{\lambda+2\mu}{2\lambda+3\mu}},$$

the maximum transverse tension will occur on the outer rim; otherwise it will occur within the cylinder. Thus a sufficiently thin annulus will have its maximum transverse tension on the outer boundary. For different values of the ratio λ/μ, the maximum traction will occur transversely or longitudinally, for a cylinder for which the ratio a/b is given. And further, for a given ratio λ/μ, the maximum traction will occur transversely or longitudinally.

Higher Approximations for Expansion of a Solid Infinite Cylinder.

The first approximation for the radial extension was of the form,

$$\rho=r+q(a_1r-r^3/8)=r+q\eta_1.$$

To obtain a second approximation we may put

$$\rho/r = 1 + q\eta_1/r + q^2\eta_2/r\,;$$

and therefore

$$d\rho/dr = 1 + q\,d\eta_1/dr + q^2 d\eta_2/dr$$

and neglect terms involving q^3. Using, again, equation (1) we obtain, if terms in q^3 are again neglected,

$$d^2\eta_2/dr^2 + d\eta_2/r\,dr - \eta_2/r^2 = -\eta_1,$$

or $$r^2 d^2\eta_2/dr^2 + r\,d\eta_2/dr - \eta_2 = -a_1 r^3 + r^5/8.$$

The solution is of the form,

$$\eta_2 = a_2 r - a_1 r^3/8 + r^5/192.$$

The radial tension T_1 is given by

$$q\left[(\lambda+2\mu)\,d\eta_1/dr + \lambda\eta_1/r\right] + q^2\left[(\lambda+2\mu)\,d\eta_2/dr + \lambda\eta_2/r\right].$$

The condition of zero radial traction on the boundary therefore yields

$$(\lambda+2\mu)\,d\eta_1/dr + \lambda\eta_1/r + q\left[(\lambda+2\mu)\,d\eta_2/dr + \lambda\eta_2/r\right] = 0$$

at $r=a$, or since the part independent of q already vanishes at $r=a$,

$$(\lambda+2\mu)\,d\eta_2/dr + \lambda\eta_2/r = 0$$

at $r=a$. Hence

$$T_1 = q(2\lambda+3\mu)(a^2-r^2)/4 + q^2\left[(2\lambda+3\mu)a_1(a^2-r^2)/8 - (3\lambda+5\mu)(a^4-r^4)/8\right],$$

and since

$$a_1 = (2\lambda+3\mu)a^2/(\lambda+\mu)$$

$$T_1 = q(2\lambda+3\mu)(a^2-r^2)/4 + q^2\left[(2\lambda+3\mu)^2 a^2(a^2-r^2)/8(\lambda+\mu) - (3\lambda+5\mu)(a^4-r^4)/8\right],$$

and, finally,

$$\eta_2/r = a_1/\left[(2\lambda+3\mu)a^2/(\lambda+\mu) - r^2/8 - \left[(3\lambda+5\mu)a^4/(\lambda+\mu) - r^4\right]/24\times 8\right.$$

$$= \frac{2\lambda+3\mu}{8(\lambda+\mu)}\,a^2\left(\frac{2\lambda+3\mu}{\lambda+\mu}\,a^2 - r^2\right) - \frac{1}{8.24}\left(\frac{3\lambda+5\mu}{\lambda+\mu}\,a^4 - r^4\right).$$

Our second approximation is therefore,

$$\begin{aligned}\rho - r &= q\eta_1 + q^2\eta_2 \\ &= q\,r\left[\frac{2\lambda+3\mu}{\lambda+\mu}a^2 - r^2\right]/8 \\ &\quad + q^2 r\left[\frac{2\lambda+3\mu}{\lambda+\mu}a^2\left(\frac{2\lambda+3\mu}{\lambda+\mu}a^2 - r^2\right)\right. \\ &\quad \left. - \frac{1}{24}\left(\frac{3\lambda+5\mu}{\lambda+\mu}a^4 - r^4\right)\right]/8.\end{aligned}$$

Any higher degree of approximation can also be obtained. Our results stand in the form,

$$\eta_1/r = {}_0a_1 + {}_1a_1 r^2,$$

$$\eta_2/r = {}_0a_2 + {}_1a_2 r^2 + {}_2a_2 r^4,$$

and, in general, the form will be

$$\eta_s/r = {}_0a_s + {}_1a_s r^2 \ldots + {}_sa_s r^{2s} \quad . \quad . \quad . \quad . \quad (2)$$

We easily see that if (2) be the form for η_s/r, then

$$\eta_{s+1}/r = {}_0a_{s+1} - \left(\frac{{}_0a_s r^2}{2\,.\,4} + \frac{{}_1a_s r^4}{4\,.\,6} \ldots + \frac{{}_sa_s r^{2s+2}}{2s+2\,.\,2s+4}\right),$$

and the condition of zero radial traction, requiring that

$$(\lambda+2\mu)\,d\eta_{s+1}/dr + \lambda\eta_{s+1}/r = 0$$

on $r=a$, yields

$${}_0a_{s+1} = \frac{{}_0a_s}{2\,.\,4}\frac{2\lambda+3\mu}{\lambda+\mu}a^2 + \frac{{}_2a_s}{4\,.\,6}\frac{3\lambda+5\mu}{\lambda+\mu}a^4 \ldots + \frac{{}_sa_s}{2s+2\,.\,2s+4}\frac{(s+1)\lambda+(2s+1)\mu}{\lambda+\mu}a^{2s},$$

so that

$$\eta_{s+1}/r = \frac{{}_0a_s}{2\,.\,4}\left(\frac{2\lambda+3\mu}{\lambda+\mu}a^2 - r^2\right) + \frac{{}_2a_s}{4\,.\,6}\left(\frac{3\lambda+5\mu}{\lambda+\mu}a^4 - r^4\right) + \ldots + \frac{{}_sa_s}{2s+2\,.\,2s+4}\left(\frac{s+1\lambda+2s+1\mu}{\lambda+\mu}a^{2s+2} - r^{2s+2}\right). \quad (3)$$

When we have obtained a solution,

$$\rho - r = q\eta_1 + q^2\eta_2 \ldots + q^s\eta_s,$$

it is therefore possible to obtain a solution to a higher degree

of approximation by remarking the relation between the coefficients in η_s/r and η_{s+1}/r and putting instead of

$$\sum_0^s {}_n a_s r^{2n},$$

as in η_s/r,

$$\eta_{s+1}/r = \sum_{n=0}^{s} \left(\frac{(n+1)\lambda + (2n+1)\mu}{\lambda+\mu} a^{2n+2} - r^{2n+2} \right) \frac{{}_n a_s}{(2n+2)(2n+4)}.$$

Rotation of an Infinite Circular Cylinder of Non-uniform Density.

We may next deal with the case of a cylinder in which the density is a function of the distance from the axis. Treating the case of the solid cylinder, we may put for the density of the cylinder when at rest,

$$\sigma = f(r) = \sum_0^N a_n r^n,$$

and N may have any value from zero (when the density is uniform) to infinity—in which case the series $\Sigma a_n r^n$ must be convergent.

The equation to be solved for a first approximation to the value of η is as before,

$$d^2\eta/dr^2 + 1/r\, d\eta/dr - \eta/r^2 = -\sigma\omega^2 r/(\lambda+2\mu).$$

The solution is evidently,

$$\eta = Ar - \frac{\omega^2}{\lambda+2\mu} \sum_0^N a_n r^{n+3}/((n+3)^2-1).$$

The fact that there must be zero radial traction on the boundary surface $r=a$, yields the condition,

$$T_1 = (\lambda+2\mu)\, d\eta/dr + \lambda\, \eta/r = 0$$

on $r=a$, giving

$$2(\lambda+\mu)A = \frac{\omega^2}{\lambda+2\mu} \sum_0^N \left[(n+3)(\lambda+2\mu) + \lambda \right] a_n a^{n+2}/(n+4)(n+2)$$

$$= \frac{\omega^2}{\lambda+2\mu} \sum_0^N \left[(n+4)\lambda + 2(n+3)\mu \right] a_n a^{n+2}/(n+4)(n+2).$$

Consequently, the value of T is

$$\frac{\omega^2}{\lambda+2\mu}\sum_0^N[(n+3)(\lambda+2\mu)+\lambda]\,(a^{n+2}-r^{n+2})a_n/(n+4)(n+2),$$

and

$$\rho-r=\frac{r\omega^2}{\lambda+2\mu}$$

$$\sum_0^N a_n\{a^{n+2}(n+4\lambda+2n+3\mu)/2(\lambda+\mu)-r^{n+2}\}/(n+2)(n+4).$$

Higher approximations to the value of η can be obtained by the method adopted in the case of uniform density.

IV. *The Straggling of α Particles by Matter.* *By* G. H. Henderson, *M.A.*, 1851 *Exhibition Scholar of Dalhousie University, Halifax, N.S.**

§ 1. *Introductory.*

When a parallel beam of α rays passes through matter, the particles gradually expend their energy in passing through the atoms of the matter, until all trace of the particles suddenly seems to vanish at the end of their range. In passing through the atoms some of the α particles lose more energy than others, so that at any point along their path some of the particles will be moving more slowly than others; also their ranges will not all be the same. The α particles may be said to be straggled out, and hence the term straggling has been applied to this phenomenon by Darwin.

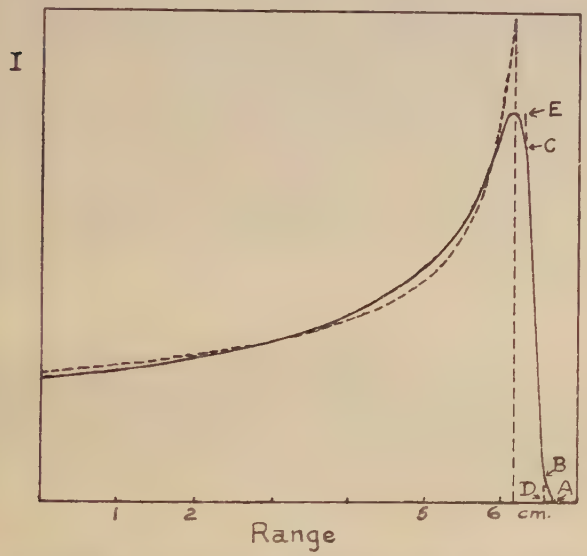
The theory of the passage of matter by α rays has been developed on the basis of the nuclear structure of the atoms of the matter, and from this theory the amount of straggling to be expected has been deduced from probability considerations. On the other hand, the straggling can be determined from experimental data in two ways.

The first method makes use of ionization data. When the ionization due to a parallel beam of α rays is measured at different points along the path of the rays, the well-known ionization curve is obtained. This curve is shown as the

* Communicated by Prof. Sir E. Rutherford, F.R.S.

full curve of fig. 1, where ionization is plotted as ordinate and distance from the radioactive source as abscissa. Now it has been shown experimentally that the (average) velocity of the α particles at any point of their path is proportional to the cube root of their remaining range. Assuming that the ionization produced is proportional to the energy lost by the α particle at any point of its path, it can at once be shown that the ionization should be inversely proportional to the cube root of the remaining range. Such a theoretical ionization curve is shown as the dotted curve of fig. 1. It

Fig. 1.



will be seen to be in approximate agreement with experiment over the first portion of the path of the α particle, but as the maximum is approached this agreement fails.

Geiger * has suggested that the ionization curve observed for a beam of α rays should be different from that of a single α particle, owing to slight variations in the ranges of the latter, *i. e.* to straggling. The ionization curve, built up of a large number of theoretical curves grouped around one of average range, will thus be modified considerably near the maximum where the ionization is changing rapidly. Hence the shape of the ionization curve near the end of the range should give an indication of the amount of straggling.

* Geiger, Proc. Roy. Soc. A. lxxxiii. p. 505 (1910).

Secondly, a more direct measure of the amount of straggling can be determined by counting the number of α particles at different points along the path of a parallel beam.

It is proposed in this paper to discuss the theoretical and experimental data on straggling, and it will be shown that the observed amount of straggling is much in excess of that allowed by theory. Further experimental evidence bearing on straggling will also be brought forward.

§ 2. *The Straggling in Air.*

It might be thought that the individual α particles are emitted with slightly different velocities, thus giving rise to straggling. It has been shown by Geiger (*loc. cit.*), however, that the α particles emitted from a thin layer of radioactive material do not differ by as much as $\frac{1}{2}$ per cent. in initial velocity. Thus the cause of the straggling must be looked for in the air itself.

As the α particle passes through the air it gives up its energy to the electrons and nuclei of the air atoms, and it is occasionally deflected through a considerable angle by close encounters with the nuclei. Different α particles will encounter different numbers and distributions of electrons and nuclei and accordingly are straggled out. The calculation of the consequent probability variations in the ranges of the individual α particles has been carried out by both Bohr * and Flamm †. They agree in showing that the nuclei produce practically no straggling. They also agree closely in the amount of straggling produced by the electrons. The straggling of various types of α rays in air, calculated by Flamm's method, is shown in the second column of Table I. The value tabulated is the distance, measured along the range, over which the number of particles in a parallel beam falls off from ·92 to ·08 of the original number. This corresponds approximately to the method of measuring the straggling from the experimental curves.

The ionization curves for three types of α rays have recently been determined with some accuracy by the writer ‡. The full curve given in fig. 1 is a reproduction of the ionization curve found for RaC. It was shown that the ionization curve from C to B (fig. 1) could be represented very approximately by a straight line. The slope of this straight line furnishes information as to the magnitude of

* Bohr, Phil. Mag. xxx. p. 581 (1915).
† Flamm, *Wien. Ber.* II *a.* cxxii. p. 1393 (1913).
‡ Henderson, Phil. Mag. xlii. p. 538 (1921).

the straggling. The easiest way of considering the matter is to imagine the straight line produced in both directions till it meets the axes of zero and maximum ionization at D and E. Then the projection of the line DE on the axis of zero ionization (or the reciprocal of its slope) is a direct measure of the straggling. The greater the straggling the greater will be the projection referred to, and as a first approximation the projection may be taken as proportional to the straggling. The values of the projections taken from the writer's curves are given in the third column of Table I. That for polonium has been determined from the curves given by Lawson *.

The curves obtained by counting the number of α particles in a parallel beam at various points along the path show that this number remains constant till near the end of the range and then falls off rapidly to zero. Most of this falling off also approximately follows a straight line. The reciprocal slope or projection of this line is a more direct measure of the straggling than the corresponding projection of the ionization curve. Measurement of the straggling by means of counting experiments is, however, very slow, as large numbers of α particles must be counted. The values obtained from the most recent and reliable counting experiments are given in the fourth column of Table I. The result for polonium is taken from a scintillation curve given by Rothensteiner †; that for RaC is from a curve obtained by Makower ‡ by photographic counting of the α particles. All the results in the Table are in millimetres and refer to air at 0° C. and 760 mm.

Table I.

Type of Rays.	Theoretical Straggling.	Straggling from Ionization Curves	Straggling from Counting Expts.
Polonium	·88	3·1	4·0
Thorium C_1	1·02	2·88	...
Radium C	1·44	2·83	4·1
Thorium C_2	1·74	2·92	...

This Table shows clearly that the observed values of the

* Lawson, *Wien. Ber.* II *a*. cxxiv. p. 637 (1915).
† Rothensteiner, *Wien. Ber.* II *a*. cxxv. p. 1237 (1915).
‡ Makower, Phil. Mag. xxxii. p. 222 (1916).

straggling are three or four times greater than those predicted by theory. Furthermore the calculated straggling increases steadily with increase of range, while that observed is constant within the limit of error. It should be pointed out that the projections given in the Table are measured as the small differences between two larger quantities, and hence are more difficult to determine with accuracy. The straggling deduced from the writer's experiments has a probable error of about 2 per cent., and it will be seen that the values for the three types of rays agree within this limit.

It was shown by the writer (*loc. cit.*) that the effect of straggling due to electronic encounters would be a tailing off of the ionization curve at the extreme end of the range. Making some simple assumptions it was shown that the calculated form of the end of the ionization curve agreed satisfactorily with the form of the curve observed between A and B (fig. 1). Thus the effect of the calculated straggling was amply accounted for by AB, leaving the much greater straggling evidenced by the straight line portion BC quite unexplained. The curves obtained by counting experiments also lead to precisely the same conclusion. In view of the failure of theory to account for this large excess straggling it is interesting to see what further information regarding it can be derived from experiment.

It is remarkable that the straggling (as measured by the projections of the ionization and also of the counting curves) should be constant for α rays differing so widely in range as those given in Table I. This can only mean that the excess straggling takes place only in the last two or three centimetres of the range. From experiments with gold foils which will be discussed later, it appears probable that the straggling is confined to the last few millimetres of the range.

Referring once more to fig. 1, it could not be expected that the straggling deduced from the ionization and the counting data would agree, for the following reasons:—The ionization curve is regarded as being built up of simple curves of different ranges grouped about a common mean. The form of the simple curve is not accurately known; the rule that the ionization is inversely proportional to the cube root of the remaining range can only be an approximation to a much more complicated law. As the shape of the simple curve cannot be taken into account, the projection of the ionization curve which is actually utilized can only give a

rough indication of the absolute magnitude of the straggling. However, comparative values of the straggling under different conditions should be given fairly accurately by the method adopted. On the other hand, in the counting experiments the assumption is made that the zinc sulphide screens or photographic plates used have the same efficiency for α rays of low speeds as for those of high speed. This assumption is not altogether justifiable.

§ 3. *Straggling in Gases other than Air.*

The ionization curve in hydrogen was determined with the same apparatus already used for air. The gas was obtained from a cylinder of compressed hydrogen stated by the makers to be of more than 98 per cent. purity. Small impurities are unlikely to affect the straggling materially.

It was again found that a considerable portion of the end of the ionization curve could be represented by a straight line. When the range of the α particles was reduced so as to give the same range as in air, the projection of the straight line was 2·05 mm. with a probable error of 3 per cent.

The straggling in air and hydrogen may be deduced from the ionization curves given by other observers. The results agree in every case within the limits of error, although the conditions for accuracy were less favourable than in the present experiments. The value 2·0 mm. for polonium in hydrogen may be obtained from some ionization curves given by Taylor *. From the results of Lawson (*loc. cit.*) for polonium the straggling was determined as 3·1 mm. in air and 2·2 mm. in hydrogen.

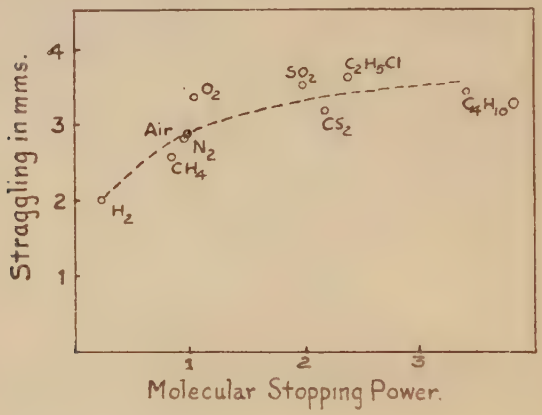
The straggling in oxygen has also been deduced from experiments made in that gas with the present apparatus, using ThC. The value found was 3·36 mm., when the range was increased to the same value as in air.

The straggling in several other gases may be deduced from the ionization curves given by Taylor (*loc. cit.*), although the error involved is probably of the order of 10 per cent. The collected results of straggling in gases are given in fig. 2, which shows the straggling plotted against molecular stopping power. The values plotted for air, hydrogen, and oxygen are from the writer's results; the remainder are taken from Taylor's curves.

* Taylor, Phil. Mag. xxi. p. 571 (1911).

Fig. 2 can only be considered to give an approximate idea of the facts, as the points are not well distributed and some may be seriously in error. It would seem, however, that the straggling increases very slowly as the stopping power of the gas is increased. It is unfortunate that the dearth of suitable gases of high stopping power makes the checking of this point difficult.

Fig. 2.



§ 4. *The Straggling due to Solids.*

The great difficulty which at once arises in determining the straggling due to solids is the uneven thickness of the solid foils used, the effect of which may completely mask the true straggling looked for. An attempt to avoid this difficulty was made by using a large number of the thinnest beaten foils of the solid obtainable; with gold, for example, as many as 128 thicknesses were used. Composite sheets of gold and other metals were placed immediately over the radioactive source (ThC) and the ionization curves determined in air with the same apparatus as before. Although a rough calculation seemed to show that the irregularities in the individual foils would be smoothed out enough to avoid masking any true increase in straggling, this result was not borne out by experiment. It was finally concluded that the increase in straggling observed was mainly, if not entirely, due to unevenness of the foils, and hence need not be gone into in detail here. In mica the increase in straggling was much the smallest, as was indeed to be expected.

Fortunately, experiments on the straggling produced by

solid foils for low velocities of the α particles gave results which were not masked by irregularities of the foils. In these experiments the foils were placed 3 mm from the middle of the ionization chamber (itself 1 mm. deep) in air at a pressure of roughly 17 cm. Reduced to air at normal pressure the distance from foil to centre of ionization chamber was therefore about ·7 mm. Most of the foils used were made up of a few thicknesses of gold leaf. The air equivalent of a single sheet was about ·45 mm. when placed directly over the source; when placed near the ionization chamber the air equivalent was about ·28 mm. The straggling of the α particles after passing through these foils was determined in the same manner as before from the ionization curve in air.

The results are given in Table II. The straggling is in millimetres, and the probable error is about 2 per cent. The third column shows the straggling observed when the foils are placed directly over the source, the steady increase with increasing number of leaves being mainly due to unevenness of foils. The fourth column shows the increase in straggling at low velocity over that at high. This increased straggling is real and almost independent of the unevenness of the foils. Results with aluminium (1·0 mm. air equivalent) and mica (8·6 mm. air eq.) are also included in the Table, but with these foils the change observed is scarcely more than the experimental error.

TABLE II.

No. of Leaves in Foil.	Straggling Foil near Ionization Chamber.	Straggling Foil near Source.	Difference.
0	2·88	2·88	...
1	3·30	2·92	·38
4	3·65	3·01	·64
12	4·00	3·28	·72
Aluminium	3·08	3·01	·07
Mica	3·16	3·06	·10

These results show quite clearly the rapid increase of straggling near the end of the range. One gold leaf nearest the ionization chamber causes nearly twice as much straggling as three leaves immediately behind it. Eight more leaves

placed behind these four again only slightly increases the straggling.

The same result was also demonstrated in a slightly different manner. A foil made up of the four gold leaves already used was placed at different distances from the ionization chamber and the straggling determined from the ionization curves as before. The results are shown in Table III. The distances given in the first column are not the actual distances from the foil to the centre of the ionization chamber, but are reduced to correspond with an ionization vessel containing air at atmospheric pressure.

TABLE III.

Distance from Ionization Chamber.	Straggling.
·7 mm.	3·65 mm.
4·7 „	3·26 „
12·9 „	2·95 „

Here again it can be seen that the straggling increases rapidly near the end of the range.

It may also be noted that the straggling (2·95 mm.) when the foil was 12·9 mm. from the ionization chamber is less than that (3·01 mm.) obtained when the foil was placed directly over the source. The difference is less than the possible error, but such a change is to be expected. For the straggling due to unevenness of the foil should be less for low velocities of the α particles on account of the decrease in air equivalent of the foil. For the same reason the differences given in the fourth column of Table II. are probably slightly too small.

From the straggling observed with different types of α particles it was pointed out in § 2 that the straggling must occur within the last two or three centimetres of the range; from the results with gold foils it seems that the straggling must be confined largely to the last few millimetres of the range.

The increase of straggling due to gold foils placed near the end of the range, though clearly marked, is small. This is quite in accordance with the view expressed in § 3 that the straggling does not increase rapidly with increase of stopping power of the substance causing the straggling. It must be remembered that the increase observed is the

increase over the straggling which would be produced by a layer of air equivalent to the gold foil used. The ionization must be measured in a gas such as air at a reasonable pressure, and hence we have the complication of the straggling due to the air between the solid and the ionization chamber, and even in the chamber itself. This sets a limitation to the amount of information to be obtained from ionization data. Accordingly the ionization experiments were not carried beyond the stage described, but it is hoped to push further the attack on the problem by more suitable methods.

An isolated experiment with iron foils may be referred to before closing this section. A sheet of iron of about 4 cm. air equivalent was placed directly over the source and the ionization curve measured when the iron was magnetized parallel with and perpendicular to the direction of travel of the α particles. The purpose of this experiment was to see if there was any change in the straggling due possibly to rearrangement or change of orientation of the electrons in the iron, a point which has been discussed by Flamm *. Alternate readings of the ionization with parallel and perpendicular fields of about 100 gauss were made at various points along the range. No appreciable difference could be detected. It should be added that the iron was very uneven in thickness, and the consequent straggling was so large (about 20 mm.) that a small change in straggling might well have been masked.

§ 5. *Summary and Conclusion.*

It has been shown in this paper that the straggling of α particles, as deduced from both ionization and counting experiments, is several times greater than that deduced from theory based on our present views of the mechanism involved in the passage of α particles through matter. It has been shown in a previous number of this Magazine that the effect of the calculated straggling can be adequately accounted for by the tailing off of the ionization curve at the extreme end of the range. The large additional straggling observed behaves quite differently. Evidence has been given in this paper to show that it increases very slowly with increase of molecular stopping power, and furthermore, that it all takes place within the last few millimetres of the range. Here we seem to be confronted with a behaviour of the α particle which present theory is unable to explain.

* Flamm, *Wien. Ber.* II *a.* cxxiv. p. 597 (1915).

Evidence obtained from the Wilson photographs also leads to the same idea, for it has been shown by Shimizu * that the observed number of ray tracks which break up into two branches near the end of the range is much greater than the number deduced from probability considerations based on our present theory of atomic structure.

It is noteworthy that this anomalous behaviour of the α particle occurs at low velocities, where practically no investigation of the scattering of α particles has been carried out on account of the experimental difficulties of dealing with slow α particles. It is at higher velocities, where the theory of scattering put forward by Sir Ernest Rutherford has been so fully verified by experiment, that the most of the theoretical straggling takes place, and this straggling has apparently been accounted for.

In conclusion I wish to express my best thanks to Professor Sir Ernest Rutherford for his kind interest and advice. I also wish to thank Mr. Crowe for the preparation of the radioactive sources.

V. *On Fluid Motion relative to a Rotating Earth.* *By* George Green, *D.Sc., Lecturer in Natural Philosophy in the University of Glasgow* †.

THE subject of this paper is at present one of considerable interest to meteorologists. Papers by the late Dr. Aitken and also by the late Lord Rayleigh on the dynamics of cyclones and anticyclones have been followed by more recent papers by Dr. Jeffreys, Sir Napier Shaw, and others. Very few actual solutions of the equations defining atmospheric motions have been obtained. In the late Lord Rayleigh's paper ‡ attention is drawn to certain general hydrodynamical principles relating to the properties of rotating fluid which can be applied to "assist our judgment when an exact analysis seems impracticable." The importance of the theorem regarding the circulation of the fluid in any closed circuit is clearly explained in its application to any actual fluid motion. In applying this theorem to fluid motion in the atmosphere, however, we must bear in mind that the motions with which we are concerned are not the actual motions of the particles in space but their motions relative to the Earth itself at each point of observation.

* Shimizu, Proc. Roy. Soc. xcix. p. 432 (1921).

† Communicated by the Author.

‡ Sc. Papers, vol. vi. p. 447.

One object of the present paper is to investigate the conditions under which the circulation theorem may be applied to atmospheric motions relative to the Earth's surface; or more generally to motions relative to any three rectangular axes which are themselves rotating about each other, with a fixed origin. In the later part of the paper one or two additional cases of motion of the atmosphere are discussed and the system of isobars corresponding to each motion determined.

In view of the problems to be considered, we shall begin by specifying the system of rotating axes most convenient in dealing with fluid motion in the neighbourhood of any point of reference O on the Earth's surface. The axis OZ is drawn upwards along the *apparent* vertical at O, and line OZ continued downwards meets the axis of the Earth at a point O′ which is taken as origin of coordinates. Then axes O′X and O′Y are drawn parallel to horizontal lines through the reference point O in directions due East and due North respectively. In the most general case to be considered the reference point O may be in motion relative to the Earth's surface, and this involves also a motion of the origin O′ if point O moves either North or South. But the motion of O′ corresponding to any moderate motion of O is very small, and for our present purpose we may regard the origin O′ as a fixed point, very near to the centre of the Earth. We shall denote by (x, y, z') the coordinates of any point referred to origin O′, and by (x, y, z) the coordinates of the same point referred to parallel axes through O. This makes $z' = z + \mathrm{R}$, where R represents approximately the radius of the Earth. The components of the velocity of any particle relative to the axes at any instant are represented by u, v, w, and the angular velocities of the axes themselves, that is, of each two axes about the third, are represented by ω_x, ω_y, ω_z, respectively. We shall introduce the particular values of $\omega_x, \omega_y, \omega_z$ corresponding to a reference point O fixed in position on the Earth, or moving relative to the Earth, when we come to deal with special problems. Referred to the above system of axes, the equations of motion of any fluid particle take the form :—

$$\frac{\mathrm{D}u}{\mathrm{D}t} + \theta_1 = -\frac{\partial \mathrm{V}}{\partial x} - \frac{1}{\rho}\cdot\frac{\partial p}{\partial x}, \quad \ldots \quad (1)$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + \theta_2 = -\frac{\partial \mathrm{V}}{\partial y} - \frac{1}{\rho}\cdot\frac{\partial p}{\partial y}, \quad \ldots \quad (2)$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} + \theta_3 = -\frac{\partial \mathrm{V}}{\partial z} - \frac{1}{\rho}\cdot\frac{\partial p}{\partial z}. \quad \ldots \quad (3)$$

In these equations, θ_1, θ_2, θ_3 represent the terms depending on the rotation of the axes themselves, being given by equations of the type

$$\theta_1 = -2\omega_z v + 2\omega_y w - \dot{\omega}_z y + \dot{\omega}_y z' + \omega_x \omega_y y + \omega_x \omega_z z' - (\omega_y^2 + \omega_z^2) x. \quad . \quad . \quad . \quad (4)$$

The function $V(x, y, z)$ represents the gravitational potential function. We have also

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}. \quad . \quad . \quad . \quad (5)$$

The equation of continuity of the fluid is then

$$\frac{D\rho}{Dt} + \rho\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) = 0. \quad . \quad . \quad . \quad (6)$$

In applying these equations we treat the atmosphere as a perfect gas in which viscosity may be neglected.

Circulation Theorem for Relative Motion.

Consider now the theorem relating to the relative circulation. We have

$$\frac{D}{Dt}(u\,dx + v\,dy + w\,dz) = \frac{Du}{Dt}dx + \frac{Dv}{Dt}dy + \frac{Dw}{Dt}dz + d\left(\tfrac{1}{2}q^2\right), \quad . \quad . \quad . \quad (7)$$

where $q^2 = u^2 + v^2 + w^2$, the square of the resultant relative velocity. By means of equations (1), (2), (3), the above equation may be rewritten in the form:

$$\frac{D}{Dt}(u\,dx + v\,dy + w\,dz) = -(\theta_1 dx + \theta_2 dy + \theta_3 dz) - \frac{dp}{\rho} - dV + d\left(\tfrac{1}{2}q^2\right). \quad (8)$$

We can now integrate each term of this equation along any curve within the fluid from any point A to any point B. This integration gives the result,

$$\frac{D}{Dt}\int_A^B (u\,dx + v\,dy + w\,dz) = -\int_A^B (\theta_1 dx + \theta_2 dy + \theta_3 dz) - \int_A \frac{dp}{\rho} - V_B + V_A + \tfrac{1}{2}q_B^2 - \tfrac{1}{2}q_A^2\ ; \quad (9)$$

and, if the integrations are applied to a closed curve

beginning and ending at the point A, we obtain

$$\frac{D}{Dt}\int_S^A (u\,dx+v\,dy+w\,dz)=-\int_S^A (\theta_1 dx+\theta_2 dy+\theta_3 dz), \quad (10)$$

where the suffix S indicates that the integration is to be taken along a definite curve S. We have assumed in obtaining (10) from (9) that V is a single valued function of (x, y, z), and that p is a function of ρ. It now appears that the rate of change of the relative circulation in any closed circuit which consists of the same fluid particles at all times is not zero unless, in addition to the above conditions, we have

$$\frac{\partial\theta_1}{\partial z}=\frac{\partial\theta_3}{\partial x}; \quad \frac{\partial\theta_2}{\partial x}=\frac{\partial\theta_1}{\partial y}; \quad \frac{\partial\theta_3}{\partial y}=\frac{\partial\theta_2}{\partial z}. \quad . \quad (11)$$

When these conditions are not fulfilled, the relative vorticity does not move with the fluid itself, and if a velocity potential exists for a certain portion of fluid at a given instant, a velocity potential will not exist for that portion of fluid at a later instant.

The first case of importance of the above conditions in relation to problems relating to the atmosphere is that in which the angular velocities ω_x, ω_y, ω_z of the axes are constants. In this case, the conditions given above take the form

$$\frac{1}{\omega_x}\delta u=\frac{1}{\omega_y}\delta v=\frac{1}{\omega_z}\delta w=\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}, \quad . \quad . \quad (12)$$

where δ represents an operator defined by

$$\delta=\left(\omega_x\frac{\partial}{\partial x}+\omega_y\frac{\partial}{\partial y}+\omega_z\frac{\partial}{\partial z}\right). \quad . \quad . \quad . \quad (13)$$

These equations have a solution of the form

$$\frac{u}{\omega_x}=\frac{v}{\omega_y}=\frac{w}{\omega_z}=f(\omega_x x+\omega_y y+\omega_z z), \quad . \quad . \quad (14)$$

where f denotes any arbitrary function. If we draw an axis to coincide with the axis defined by the resultant of the three component rotations ω_x, ω_y, ω_z, then $(\omega_x x+\omega_y y+\omega_z z)$ is equal to $\Omega R\cos\phi$, where Ω is the resultant of $(\omega_x, \omega_y, \omega_z)$ and R is the line joining the origin to the point x, y, z. That is, u, v, w are functions of p the perpendicular from the origin to a plane through the point (x, y, z) perpendicular to the axis of the resultant rotation.

When the fluid is incompressible, and when a compressible fluid is moving in such a way that $\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}$ is zero, a solution of a different type obtains. The solution in this case may be written in the form

$$\left.\begin{aligned} u&=f_1\{(\omega_y x-\omega_x y),\ (\omega_z x-\omega_x z)\},\\ v&=f_2\{(\omega_y x-\omega_x y),\ (\omega_z x-\omega_x z)\},\\ w&=f_3\{(\omega_y x-\omega_x y),\ (\omega_z x-\omega_x z)\},\end{aligned}\right\}\quad .\quad .\quad .\quad (15)$$

where f_1, f_2, f_3 are arbitrary functions subject only to the condition $\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0$. This solution includes as a particular case any motion of rotation of the atmosphere as a solid about the axis of the Earth.

The solutions which we have above obtained make it clear that the fluid motions relative to rotating axes in which the relative circulation moves with the fluid belong to a very restricted type. A relative motion, for instance, similar to that taking place in a free vortex, does not fulfil the conditions required for permanence of the velocity potential, and therefore no steady motion of this type could take place in the atmosphere—as has been assumed to be the case.

The conditions which we have found to be necessary for the validity of the circulation theorem when the fluid motion is relative to rotating axes, may be obtained in a manner different from that employed above. Taking Ξ, H, Z to denote the components of angular velocity of a fluid element, and U, V, W to denote components of linear velocity of the element, each referred to fixed axes which coincide at instant t with the instantaneous positions of the moving axes, we may derive the conditions from the equations employed by von Helmholtz in his papers on vortex motion :—

$$\frac{\mathrm{D}}{\mathrm{D}t}\Xi=\Xi\frac{\partial \mathrm{U}}{\partial x}+\mathrm{H}\frac{\partial \mathrm{U}}{\partial y}+\mathrm{Z}\frac{\partial \mathrm{U}}{\partial z}-\Xi\left(\frac{\partial \mathrm{U}}{\partial x}+\frac{\partial \mathrm{V}}{\partial y}+\frac{\partial \mathrm{W}}{\partial z}\right),\quad (16)$$

with two other similar equations. With ξ, η, ζ to represent the components of relative angular velocity of an element of fluid, referred to the moving axes, we have,

$$\Xi=\xi+\omega_x\ ;\quad \mathrm{H}=\eta+\omega_y\ ;\quad \mathrm{Z}=\zeta+\omega_z\ ;$$

$$\mathrm{U}=u-\omega_z y+\omega_y z\ ;\quad \mathrm{V}=v-\omega_x z+\omega_z x\ ;\quad \mathrm{W}=w-\omega_y x+\omega_x y\ ;$$

and

$$\frac{\mathrm{D}}{\mathrm{D}t}\Xi=\frac{\mathrm{D}\xi}{\mathrm{D}t}-\omega_z\eta+\omega_y\zeta.$$

By means of these relations we can readily transform (16) and obtain the corresponding equations for the rates of change of the circulation components of an element of fluid referred to the rotating axes ; in this way we find

$$\frac{D\xi}{Dt} = \xi\frac{\partial u}{\partial x} + \eta\frac{\partial u}{\partial y} + \zeta\frac{\partial u}{\partial z} - \xi\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)$$
$$+ \omega_x\frac{\partial u}{\partial x} + \omega_y\frac{\partial u}{\partial y} + \omega_z\frac{\partial u}{\partial z} - \omega_x\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right), \quad (17)$$

with the corresponding equations in η and ζ. Now the hydrodynamical theorem that relates to the permanence of a velocity potential for the motion of a given portion of fluid and the theorem of the permanence of the circulation of an element of fluid depend on equations (16). The equations which we have obtained for the relative circulations reduce to these equations exactly when the conditions expressed in (12) are fulfilled ; and these conditions must accordingly be fulfilled in order that the theorems referred to may apply to the relative motion, in the same way as they apply to the actual motion.

Particular Cases of Motion Relative to the Earth.

We shall now discuss one or two particular cases of fluid motion relative to the Earth, and we shall, to begin with, take the reference point O as a fixed point on the surface of the Earth at latitude ϕ degrees North. The angular velocity components $\omega_x, \omega_y, \omega_z$ have then the values 0, $\Omega\cos\phi$, $\Omega\sin\phi$, respectively, where Ω represents the angular velocity of rotation of the Earth about its axis. If we now let ϖ represent the perpendicular distance from any point x, y, z to the axis of the Earth, we can write the equations of motion (1), (2), (3), in the form

$$\frac{Du}{Dt} - 2\Omega\sin\phi\,.\,v + 2\Omega\cos\phi\,.\,w = -\frac{\partial V'}{\partial x} - \frac{1}{\rho}\cdot\frac{\partial p}{\partial x}, \quad (18)$$

$$\frac{Dv}{Dt} + 2\Omega\sin\phi\,.\,u = -\frac{\partial V'}{\partial y} - \frac{1}{\rho}\cdot\frac{\partial p}{\partial y}, \quad (19)$$

$$\frac{Dw}{Dt} - 2\Omega\cos\phi\,.\,u = -\frac{\partial V'}{\partial z} - \frac{1}{\rho}\cdot\frac{\partial p}{\partial z}, \quad (20)$$

where $V' = V - \frac{1}{2}\Omega^2\varpi^2$. V' is, in fact, the potential function

corresponding to apparent gravity, so that $-\frac{\partial V'}{\partial z}$ at point ($x=0$, $y=0$, $z=0$) is the value of $-g$ at the reference point O. In applying the above equations to motion of the atmosphere, we may take

$$\frac{\partial V'}{\partial x}=0, \quad \frac{\partial V'}{\partial y}=0, \quad \frac{\partial V'}{\partial z}=g, \quad . \quad . \quad (21)$$

in the immediate neighbourhood of O in the region within which the value of apparent gravity may be regarded as constant in direction and amount. If we neglect a change of direction of one degree in g, our equations (18) to (21) would then represent conditions of motion within a radius of about seventy miles from point O. In order to render our equations suitable to represent circulations of air of diameter exceeding, say, 150 miles, we might employ the approximate values

$$\frac{\partial V'}{\partial x}=\frac{gx}{R}; \quad \frac{\partial V'}{\partial y}=\frac{gy}{R}; \quad \frac{\partial V'}{\partial z}=g, \quad . \quad . \quad (22)$$

wherein we neglect the variation of g with height.

If we exclude certain cases of motion relating specially to the tides, very few solutions of the above equations have been recorded. In order to make ourselves familiar with the types of fluid motion possible in the atmosphere it is of interest to examine all solutions which can be obtained having a bearing on meteorological problems. We, accordingly, take first the steady rotational motions of incompressible fluid under the force of gravity alone.

We may take the boundary condition $w=0$ to apply at the surface of the Earth. A simple rotation of all the fluid about a vertical axis through O, with a uniform angular velocity ω, would be represented by

$$u=-\omega y; \quad v=+\omega x; \quad w=0. \quad . \quad . \quad (23)$$

With these values in equations (18) to (20), it would be impossible to satisfy (19) and (20) simultaneously; but the motion represented by

$$u=-\omega(y-\beta z); \quad v=\omega x; \quad w=0, \quad . \quad . \quad (24)$$

fulfils all the conditions contained in our equations, provided β is given by

$$\beta=\frac{2\Omega\cos\phi}{\omega+2\Omega\sin\phi}, \quad . \quad . \quad . \quad . \quad . \quad (25)$$

and the pressure p is given by

$$\frac{p}{\rho} = -V' + \tfrac{1}{2}(\omega^2 + 2\omega\Omega \sin\phi)\{x^2 + (y-\beta z)^2\} + \frac{p_0}{\rho}. \quad (26)$$

The approximate value of V' is $+g\left(\frac{x^2+y^2}{2R} + z\right)$, and p_0 is the pressure at the reference point O. In the motion indicated above each particle of fluid moves in a horizontal circle whose centre lies in the line $(y=\beta z,\ x=0)$. This line lies in the meridian plane through point O and is inclined to the vertical at O at an angle θ towards the North, where $\tan\theta=\beta$. When ω is very small in relation to Ω this inclined axis is almost parallel to the axis of the Earth; and when ω is large this axis comes almost to coincidence with the apparent vertical at O. With ω very large the motion described above corresponds very closely with a uniform rotation of the fluid about a vertical axis, as in the case of a simple forced vortex.

Another case of motion of incompressible fluid of interest in the same connexion is that represented by

$$u = -\omega y\ ; \qquad v = \omega x\ ; \qquad w = 2\Omega \cos\phi\ .\ x, \quad . \quad (27)$$

with

$$\frac{p}{\rho} = -V' + \tfrac{1}{2}(\omega^2 + 2\omega\Omega \sin\phi)(x^2+y^2) + 2\Omega^2 \cos^2\phi\ .\ x^2 + \frac{p_0}{\rho}, \quad . \quad . \quad . \quad (28)$$

as the equation showing the distribution of pressure. This motion differs from that first discussed in not being exactly horizontal. The plane of motion of each particle of fluid passes through the line OX, and is inclined to the horizontal plane XOY at an angle θ given by $\tan\theta = 2\Omega\cos\phi/\omega$. When the angular velocity of rotation ω is very large compared with $\Omega\cos\phi$, the plane of motion of each particle is practically horizontal, and the motion then corresponds very closely with that of simple rotation of all the fluid as a solid about a vertical axis. When ω becomes small, on the other hand, the inclination of the plane of motion of each particle of fluid to the horizontal increases. The two motions, represented by (24) and (27) respectively, are almost identical when ω is very large, and they differ entirely when ω is very small. It would be interesting to investigate the manner in which a fluid, such as water, subsides to rest from an initial condition of steady rotation about a vertical axis. The solution represented by (24), (26), would appear to be the exact solution for steady rotation of water in small scale experiments.

The motions considered above are motions of any incompressible fluid and do not indicate, except as approximations, conditions of motion possible in the atmosphere. One or two solutions of a similar type can be obtained which refer to incompressible fluid and accordingly represent motions possible in the atmosphere. Consider now the motion represented by

$$u=-\omega(y-\beta z)\,; \quad v=0\,; \quad w=0. \quad . \quad . \quad (29)$$

The equations of motion of any element of fluid in this case are

$$0=-\frac{gx}{\mathrm{R}}-k\frac{\partial}{\partial x}\log\rho, \quad . \quad (30)$$

$$-2\Omega\sin\phi\,.\,\omega(y-\beta z)=-\frac{gy}{\mathrm{R}}-k\frac{\partial}{\partial y}\log\rho, \quad . \quad (31)$$

$$2\Omega\cos\phi\,\omega(y-\beta z)=-\ g-k\frac{\partial}{\partial z}\log\rho, \quad . \quad (32)$$

and these equations are satisfied provided the pressure system throughout the fluid is that indicated by

$$k\log\rho=\Omega\sin\phi\,.\,\omega(y-\beta z)^2-g\left(\frac{x^2+y^2}{2\mathrm{R}}+z\right)+k\log\rho_0. \quad (33)$$

where $\beta=\cot.\phi$, and ρ_0 is the density of the fluid at the reference point O. The continuity equation is also satisfied provided we can neglect the term $(g\omega/\mathrm{R})x(y-\beta z)$. This condition limits considerably the extent of the region around O to which our solution is applicable, as stated earlier. Within the region to which the above applies the isobars at the surface of the Earth run East and West, being determined by

$$k\log\rho=\Omega\sin\phi\,.\,\omega y^2+k\log\rho_0, \quad . \quad . \quad . \quad (34)$$

which indicates a system symmetrical on the two sides of the East and West line drawn through reference point O. In this case the isobars become closer as we proceed North or South from point O. They are also parallel to the lines of flow of the fluid.

The coefficient $\frac{g}{\mathrm{R}}$ has the value $1{\cdot}5\times10^{-6}$ with the foot and the second as units; while $2\Omega\sin\phi$ has the value $1{\cdot}03\times10^{-4}$ at latitude 45°. Certain cases of interest arise in which the terms containing $\frac{g}{\mathrm{R}}$ may be neglected. For

example, we may take the case of a uniform east or west wind over a considerable region. In this case

$$u=c, \quad v=0, \quad w=0, \quad . \quad . \quad . \quad . \quad (35)$$

and the pressure distribution consistent with this motion is represented by

$$k \log \rho = 2\Omega c \, (\cos \phi \, z - \sin \phi \, y) - gz + k \log \rho_0, \quad . \quad (36)$$

where ρ_0 refers to the density of air at the reference point O. The isobars at the Earth's surface are in this case a uniform system running due East and West.

A similar case is that given by

$$u=0, \quad v=cx+d, \quad w=0, \quad . \quad . \quad . \quad (37)$$

which represents a wind towards the North, while the corresponding pressure distribution is that represented by

$$k \log \rho = \Omega \sin \phi \, (cx^2 + 2dx) - gz + k \log \rho_0. \quad . \quad (38)$$

The isobars are again a system of straight lines, but running north and south, and uniformly spaced when $c=0$.

In a similar manner we find that

$$u=c_1, \quad v=c_2, \quad w=0, \quad . \quad . \quad . \quad . \quad (39)$$

corresponds to a system of straight isobars represented by

$$k \log \rho = 2\Omega \sin \phi \, (c_2 x - c_1 y) + 2\Omega \cos \phi \, . c_1 z - gz + k \log \rho_0. \quad (40)$$

The isobars are again lines of flow of the air, as in each case considered above.

The case of motion corresponding most closely to a cyclonic or anticyclonic circulation is that discussed in an earlier paper *, represented by

$$u = -\omega(y - \beta z) \; ; \quad v = \omega x \; ; \quad w = 0. \quad . \quad . \quad (41)$$

In this case

$$\beta = \frac{2\Omega \cos \phi}{\omega + 2\Omega \sin \phi}, \quad . \quad . \quad . \quad . \quad . \quad (42)$$

and the pressure distribution is that represented by

$$k \log \rho = \tfrac{1}{2}(\omega^2 + 2\omega\Omega \sin \phi) \, \{x^2 + (y - \beta z)^2\} - gz + k \log \rho_0. \quad . \quad . \quad . \quad (43)$$

The term $\frac{g}{R} \omega x z$ must be small in order that the continuity equation may be fulfilled.

* Phil. Mag. vol. xli. April 1921 ; vol. xlii. July 1921.

In each case considered we have only to replace Ω by an increased value Ω' to obtain a motion for which the system of isobars travels eastward at a uniform speed ($\Omega' > \Omega$).

Each of the above solutions has been given to apply to an isothermal atmosphere, and in every case considered the fluid moves so that no element of fluid undergoes change of density. That is, $\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0$ at each point. Provided this condition is fulfilled, any solution obtained for motion of an atmosphere, all at one temperature, can readily be transformed to suit an atmosphere in convective equilibrium ($p = k\rho^\gamma$), or one in which pressure is any given function of density. Thus, taking the motion represented by (41) above in an atmosphere in convective equilibrium, we have merely to replace $\log \rho$ and $\log \rho_0$ in (43) by $\frac{\gamma}{\gamma - 1}\rho^{\gamma - 1}$ and $\frac{\gamma}{\gamma - 1}\rho_0^{\gamma - 1}$ respectively, all other conditions being unchanged.

VI. *The Breaking Stress of Crystals of Rock-Salt.* *By* Prof. G. N. Antonoff, *D.Sc.(Manch.)* *.

In a paper published in Phil. Mag. vol. xxxvi. Nov. 1918, I have developed a theory of surface tension under the assumption that the attraction of molecules is due to electrical or magnetic forces, or both. Instead of assuming a uniform field round the molecules as it is generally accepted according to Laplace, I accepted the view that the molecules act as electrical doublets, and from the theory of potential I deduced that the attraction between them must be inversely proportional to the 4th power of the distance, provided the distance between the doublets is large compared with their respective lengths.

It was shown that the attraction between the doublets can be represented by an expression of the type

$$\frac{kl^2}{d^4},$$

where k is a constant, l the length of the doublet, and d the distance between them. In these calculations the magnetic forces were disregarded altogether, as the law of attraction between small magnets would be just the same, so that they could only have an effect on the value of k.

* Communicated by Dr. J. W. Nicholson, F.R.S.

For the surface tension, the expression was given as

$$\frac{kl^2}{d^4}\Big/d,$$

or assuming $\frac{1}{d^3}=p$ where p is the number of molecules per unit volume, the expression for the surface tension α becomes

$$\alpha=kl^2p^{5/3};$$

instead of p we may put $\frac{\delta_1-\delta_2}{M}$, where δ_1 is the density of the liquid, δ_2 that of the saturated vapour, and M the molecular weight.

Thus
$$\alpha=kl^2\left(\frac{\delta_1-\delta_2}{M}\right)^{5/3}. \quad . \quad . \quad . \quad . \quad . \quad (1)$$

It was also shown that the internal pressure P can be calculated by the formula

$$P=2\alpha p^{1/3} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

or
$$P=2\alpha\left(\frac{\delta_1-\delta_2}{M}\right)^{1/3}. \quad . \quad . \quad . \quad (3)$$

In other words, the intrinsic pressure can be calculated from the surface tension if the molecular weight of the liquid is known. Thus for normal or non-associated liquids it should be possible to calculate the normal pressure from the value of the surface tension.

It should be pointed out that the assumption made in the above theory, that the length of the doublet is small compared with the intramolecular distance, is not necessarily the right one.

In the expression (2) this law is, however, eliminated, and the same expression is obtained for any law of molecular attraction. The figures for P obtained from the above expression agree with those from indirect evidence. However, experimentally it is not possible to determine P directly, owing to the mobility of the particles of liquids which always adjust themselves so that the molecular pressure is inappreciable.

It is not so in the solid or crystalline state, in which the particles have a definite orientation, and where the internal pressure can be determined by a direct experiment. It is sufficient to apply to a crystalline body such a weight as would overcome the attraction of the molecular forces and cause the disruption of the body. The force applied is not

necessarily the same in all directions, and it is therefore necessary to specify the direction in which it is to be applied.

The question arises now whether it is possible to calculate the surface tension of a solid body.

For the solid state there is no direct method of determining the surface tension, all methods used for liquids being inapplicable in this case. Some attempts were made to estimate the surface tension of solids from indirect evidence. For example, Ostwald * and Hulett † calculated the surface tension of some calcium and barium salts on a basis of a certain theory from the solubility data. The figure given for the latter is about 4000 dynes per cm. From the point of view of our theory, it seems possible to calculate the surface tension by the use of formula (2) by determining experimentally the internal pressure per square cm. of the cross-section, if the molecular structure of the substance in the crystalline state is known.

At the present time the X-ray analysis throws a light on the above question.

For example, according to W. H. and W. L. Bragg, the crystal of rock-salt consists of charged ions situated at regular distances from one another.

Such a case is somewhat different from the one discussed in my paper (*loc. cit.*). Here it is necessary to assume that $l=d$, where l is the length of the doublet and d the distance between them, under which conditions the ordinary inverse square law must hold true. The attraction between the charges in a row is equal to

$$\frac{e^2k}{d^2},$$

where e is the elementary charge, and the value of k is the sum of a series $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\ldots=0{\cdot}6931$.

Assuming that the adjacent rows have no effect upon the charges, the expression for the surface tension is of the form

$$\frac{e^2k}{d^2}\Big/d=e^2kp,$$

where $p=$number of particles per unit volume.

For the normal pressure the expression will be

$$P=ke^2p^{4/3}.$$

* *Zeit. Phys. Chem.* xxxiv. p. 503 (1900).
† *Zeit. Phys. Chem.* xxxvii. p. 385 (1901).

For rock-salt

the density $=2{\cdot}15$,

$$M = 1{\cdot}64 \times 58{\cdot}5 \times 10^{-28},$$

$$p = \frac{\delta}{M} = 2{\cdot}24 \times 10^{22}.$$

Assuming that k is approximately $=0{\cdot}7$, the tension is

$$e^2 \times 0{\cdot}7 \times 2{\cdot}24 \times 10^{22} = 22{\cdot}1 \times 10^{-20} \times 0{\cdot}7 \times 2{\cdot}24 \times 10^{22}$$
$$= 3500 \text{ dynes per cm. (approximately).}$$

The figure obtained is of the same order of magnitude as figures derived by Ostwald and Hulett (*loc. cit.*) for barium salts.

The normal pressure P would be accordingly

$$P = 98{\cdot}7 \times 10^9 \text{ dynes per square cm.}$$

It is interesting to see now how far the above results agree with the experimental evidence.

An experiment was performed as follows :—

I took a good specimen of rock-salt crystal and I cut a prism of the section about 15 square mm. and about 2–3 cm. long. I used a suitable cement to hold the piece from both ends, and by applying a suitable weight produced a rupture of the crystal into two halves. Measuring the cross-section of the rupture accurately, I calculated the weight required to produce the rupture per square cm. I have repeated the experiment many times with different samples of rock-salt. If the crystal is well formed, the agreement between individual experiments is fairly good. In one series of experiments, I cut the prisms so as to have several samples cut parallel to the three principal axes. I have done the experiment with three pieces for each direction. For one direction I obtained :

89 lb. per cm.
103 „ „
83 „ „

In the other two directions the results were identical.

In some cases it happened that the rupture took place under a much smaller weight. This, however, could be attributed either to some faults in the structure of the crystal, or to some other disturbances. Such measurements were simply disregarded.

Taking as the average value 91·7 lb., or 41·5 kgrs. per square cm., one can calculate the inward pull per row of unit length.

This is obtained by dividing P by $p^{1/3}$, and becomes

$$\frac{P}{p^{1/3}} = \frac{415000 \times 981}{2{\cdot}82 \times 10^7} = 1{\cdot}4 \text{ dynes per cm.}$$

If the attraction between the charges in a row is not appreciably influenced by the adjacent rows, this value will represent the accurate value of surface tension in the direction coinciding with the vertical axis. If the field is symmetrical in all three directions parallel to the main axis, this figure will characterize the surface tension of rock-salt in all three directions.

The symmetrical structure indicated by W. H. and W. L. Bragg is in accord with the experiment in this sense, but on the other hand, the above figure is about 1000 times less than one would expect.

One could expect such a small value if the salt consisted of molecules with a very small polarity situated at large distances from one another. But in such a case the force in the direction coinciding with that of the doublet would have to be twice that in the perpendicular direction. This, however, is not the case.

In my paper (*loc. cit.*) it was shown that in a case of small doublets the adjacent rows have practically no influence on the attraction between the doublets.

However, in a case of charges situated at regular distances from one another, such seems not to be the case.

If the charges influence one another, one can expect the forces to be weaker in the middle of the substance, and much bigger at the surface where the above effect is only one-sided.

It is therefore probable that the above figure 1·4 dynes per cm., although quite characteristic for the substance, is not the actual value of the force in the surface layer. The calculation of these effects is not easy, owing to the fact that one has to deal with a very slowly converging series. I satisfied myself that these influences may be appreciable, but I do not see clearly at the present time whether they can account for the weakening of the forces about 1000 times, or even more.

6 Featherstone Buildings,
High Holborn,
London, W.C. 1.

VII. *The Measurement of Chance.*

By Norman Campbell, *Sc.D.**

Summary.

It is maintained that the chance of an event happening is always a physical property of a system, measured by a process of derived measurement involving the two fundamental magnitudes—number of events and number of trials.

Chances are not measurable by a process of fundamental measurement. But the calculation of chances is analogous to fundamental measurement. It is usually theoretical, and is valuable only in so far as the calculated chances are confirmed by measurement.

When a proposition concerns a system characterized by a chance, it may sometimes (but by no means always) be regarded as having a definite probability determined by that chance. The probability of propositions which do not concern systems characterized by chances has nothing to do with chance.

1. It is generally recognized that there are two kinds of "probability." There is (1) the probability (of the happening) of events, and (2) the probability (of the truth) of propositions. Etymologically the term belongs more properly to the kind second of probability, and it will be confined to that kind in this paper. For the first kind, the term "chance," often used in some connexions as a synonym of probability, is available. Accordingly we shall speak throughout of the chance of an event happening and of the probability of a proposition being true.

Various opinions have been entertained concerning the relation between chance and probability and between the methods of measuring them. Some have held that chance, some that probability, is the more fundamental conception, and that the measurement of the less fundamental depends on that of the more fundamental conception. Others have held that only one of the two, or neither, is measurable. The conclusion towards which this paper is directed is that chance, in the sense primarily important to physics, is a physical property measurable by ordinary physical measurement. This view is similar to that held by Venn; indeed, it is

* Communicated by the Author.

probably the view that Venn would have held if he had ever considered the nature of physical measurement. But the further view often attributed to Venn, though it is doubtful whether he actually held it, that probability is always measurable in terms of chance—this view will not be upheld, but, so far as it is discussed at all, will be combated.

Many of the ideas and terms used in the discussion are explained more fully in my 'Physics,' to which references are made by the letter P. In fact, this paper may be regarded as a substitute for pp. 168–183 of that book, some of the difficulties of which are avoided by the alternative method of treatment. However, I should like to add that I do not accept any of the criticisms that have been directed against those pages by others.

2. Suppose we are presented with a pair of dice, and asked what is the chance that when one of them is thrown it will turn up six. The answer may be different for the two dice. If one of them is accurately cubical with its centre of mass accurately at the centre of the cube, while the other has corners and edges variously rounded and is loaded so that the centre of mass is appreciably nearer one face than another, then the answer will be different. On the other hand, if in all respects the dice are the same, then the answer will be the same; even if they are both inaccurate in form and both loaded—the inaccuracy of form and the loading being the same—the chance that they will turn up six will be the same. This chance is something uniformly associated with and changing with the structure of the die, just as is (say) the electrical resistance. This uniform association of the resistance with the other characteristics of the die is what we assert when we say that the resistance is a physical property of the die, and accordingly the chance of turning up six is a physical property as much as the resistance.

Moreover, the chance is measured by essentially the same process as that by which the resistance is measured. Resistance is measured (in its original meaning) as a derived magnitude by means of a numerical law (P. Ch. xiii.). We place two electrodes in contact with opposite faces of the die, and measure the current which flows through it when measured potential differences are maintained between the electrodes. We then plot current against P.D., and find we can draw a straight line through (or more accurately among) the resulting points. The fact that the graph is a

straight line passing through the origin shows that a numerical law of a certain form holds, and therefore that the die is characterized by a single definite magnitude, which is what we mean by resistance ; the slope of the line tells us the numerical value of this magnitude. When we proceed to measure the chance of turning up six, we make several groups of trials, measure in each group the number of trials and the number of those in which six turns up. We plot these two fundamental magnitudes against each other, and find that a straight line can be drawn through (or among) the points. The fact that the graph is straight and passes through the origin tells us that the die is characterized by a definite magnitude, which is what we mean by chance ; the slope of the line tells us the numerical value of this magnitude.

3. The resemblance is exact in all essentials. But as the conclusion that chance is an ordinary physical magnitude does not seem to be universally accepted, some objections may be considered.

The first may be (though I am not sure that it will be) raised by those who denounce the "frequency theory" of probability. They might say that, though the derived magnitude, estimated in the manner described, is a true or approximate measure of the chance, yet it is not what is meant by the chance—that is something much more abstruse. Such an objection can only be met by stating more clearly what is asserted, and recognizing any difference of opinion that remains as insoluble. What I assert is (1) that all chances determined by experiment are determined by a relation between frequencies, and (2) that chances are important for physics only in so far as they represent relations between frequencies. Few examples can be cited in support of (1), for chance in physics is usually a theoretical and not an experimental conception ; but it may be suggested that anyone who proposed to attribute to the chance of a given deflexion of an α-ray in passing through a given film any value other than that determined by frequency, could convince us of nothing but his ignorance of physics. In support of (2) it may be pointed out that the chance, which is such an important conception in the statistical theories of physics, enters into the laws predicted by those theories only because it represents a relative frequency.

4. A second objection may be based on the fact that the straight line has to be drawn among and not through the experimental points. It may be readily admitted that this fact shows that the chance cannot be estimated with perfect accuracy. But there is also some uncertainty in determining the resistance ; and since I am concerned only to enforce the analogy between chance and resistance, the admission is innocuous. If it is urged that this uncertainty shows that the derived magnitude cannot be the chance, because chance is something to which a numeral may be attached with mathematical accuracy, then it is replied (as in answering the first objection) that such a chance, to which no experimental error is attached, is something totally irrelevant to physics.

But the objection may be put in a less crude form. It may be urged that, in the matter of experimental error, there is a fundamental difference between resistance and chance. For in the latter, but not in the former, the error is something essential to the magnitude ; we can conceive of a resistance measured without error, but not of a chance measured without error ; if all the points lay accurately on the line, then the magnitude measured by its slope would not be a chance. Again, there is a simple relation between the average error about a point on the "chance" line and the co-ordinates of that point ; while in the "resistance" line the relation is much more complex, and depends on the exact method of measuring the current and potential. All this is quite true, and would be important if we were considering the theory of chance or of resistance. There is a great difference in those theories; we suppose that the "real" Ohm's law holds between the real and not the measured magnitudes of the current and potential, while there is no real magnitude involved in the chance relation. But we are not considering theory but experiment ; I am only asserting that chance is an experimentally measured magnitude. The fact that the errors in the two cases are differently explained does not affect the fact that there are errors in both cases, and that the problem of determining the derived magnitude in spite of these errors is precisely the same.

5. As a third objection it might be urged that the two measurements are not really similar, because the chance is not really determined by the slope of the line, but by the ratio of the two numbers when they are sufficiently great. Here is a misconception which it is important to correct. If we know that the happening of the events is determined

by chance, then it is true that we need only plot one point on the line ; and the distribution of the errors is such that the relative error of a determination from a single point is less the greater the number of trials involved. We shall group all our observations together, so as to make their total number as great as possible. But similarly, if we know that the material of the die obeys Ohm's law, one observation is sufficient to determine its resistance ; and the accuracy of the determination will be greatest if we choose the measuring current within a certain range. An even closer parallel would be obtained if we took in place of resistance the derived magnitude, uniform velocity. If we knew that the velocity was uniform, we should choose our time and distance as great as possible, and determine the velocity from this single pair of values without troubling to plot smaller values.

But in order that determination by a single point should be legitimate, we must know that the events really are determined by chance, and the only test of chance is that, when a series of points are plotted in the manner described, the only regularity discoverable in them is that they lie about the straight line. Their distribution about that line must be random. Thus, to take Poincaré's excellent example, if the trials were made by selecting the first figure of the numerals in a table of logarithms in the conventional order, and the events were the occurrence of the figure 1, the plotted points would lie on the whole about a line with a slope of 1/10. But a regularity of distribution about that line would be apparent ; we should have a series of points all lying above the line followed by a set all lying below. If, on the other hand, we took the last figure of the numerals, no such regularity would be apparent ; the distribution of the points about the line would be random ; the events would be dictated by chance.

It is of the first importance to insist that in measuring a chance we are picking out the only regularity that we can find in some sequence of phenomena, leaving a residuum which is purely random. Randomness is a primary conception, incapable of further definition ; it cannot be explained to anyone who does not possess it. It is based, I believe, on observation of the actions of beings acting consciously under free volition ; and it is subjective in the sense that what is random to one person may not be random to another with fuller knowledge (P. p. 203). There are certain forms of distributions that are random to everybody ; it is this common randomness, objective in the sense in which all the

subject-matter of science is objective, that is the characteristic of the objective chance which is physically measurable. Chance is applicable only to events which contain an element which is wholly and completely random to everybody *.

6. We shall then base our further discussion on the assumption that any physically significant chance (of the happening of an event) is a measurable derived magnitude, a property of the system concerned in that event, determined by a linear numerical law relating the fundamental magnitudes, number of events and number of trials. It is thereby implied that the "errors" from the law are random, for otherwise the law would not be linear. The definition of chance as the limiting ratio of the fundamental magnitudes as they tend to infinity is identical with that given, if it is known as an experimental fact that the magnitudes of the errors fulfil certain conditions which need not be discussed in detail here; these conditions are not inconsistent with the randomness of the errors.

Chance as a fundamental magnitude.

7. Another important question may be raised, again suggested by the analogy with resistance. Resistance means the derived magnitude defined by Ohm's law. But actually resistance is measured nowadays, not as a derived, but as a fundamental magnitude, in virtue of the Kirchhoff laws for the combination of resistances in series and parallel †. Can chance, though meaning the derived magnitude, be measured independently as fundamental?

In order that a property may be measured as a fundamental magnitude, it is necessary that satisfactory definitions of equality and of addition should be found (P. Ch. x.). In addition, some numerical value must be assigned arbitrarily to some one property, which with all others can

* In this sense the last figure of the logarithm is not wholly dictated by chance; for we know that there must be some regularity in the distribution of the points about the straight line, even if we cannot say exactly what it is. In the strictest sense, therefore, there is no such thing as the chance of the last figure being 1. But there are events which are, at present at least, wholly dictated by chance in this sense, *e. g.* the distintegrations of a radioactive atom. Here I do not think anyone has imagined what kind of regularity there can be, except the falling of the plotted points about the straight line which determines the chance.

† Ultimately measured, that is to say, by the makers who calibrate our resistance boxes. In the laboratory we use a method which is essentially that of judgment of equality with a graduated instrument.

be compared by means of these definitions. Since we take the meaning of chance to be that of the derived magnitude, the definitions will be satisfactory if they are in accordance with the derived process of measurement; but we shall not succeed in establishing an independent system of fundamental measurement, unless the definitions are such that they can be applied without resort to that process.

The arbitrary assignment is usually made by attributing the value 1 to the chance of an event which always happens as the result of a trial. The only question that can arise here, namely whether all other chances can be connected with this chance by addition and equality, will be considered presently. The definition of addition presents no difficulty. The chance of A happening is the sum of the chances of x, y, z, ... happening if, x, y, z, ... being mutually exclusive alternatives, A is the event which consists in the happening of either x or y, or z, ... This proposition is introduced in all discussions of chance, but it is often introduced as a deduction and not as a definition. The inconsistencies which result from such a procedure are discussed in P. pp. 174, 184, 185. As a definition, it is satisfactory in our sense, for measurements by the derived method would show that, in such conditions, the chance of A is the sum of the chances of x, y, z, ..., and yet it does not presuppose such derived measurements. If the points in the derived measurement lay on the straight line, this result would be a direct consequence of the definition of the derived magnitude; but since they do not, it can be deduced from that definition only if some assumption about the distribution of the errors is made. The assumption that the errors are random would probably suffice if randomness could be strictly defined; since it cannot, the agreement of the proposed definition of addition with the results of the derived process of measurement must be regarded as an experimental fact. The definition is thus precisely analogous to that used in the fundamental measurement of resistance, namely that resistances are added when the bodies are placed in series.

The definition of equality is much more difficult; in fact, it is the stumbling block of many expositions of the measurement of chance. For resistances we can say that bodies are equal if, when one is substituted for another in any circuit, the current and potentials in that circuit are unchanged; that definition does not involve a knowledge of Ohm's law and of the derived measurement. The only attempt at an analogous definition for chance, of which I am aware, is that based on the principle of sufficient reason; chances

are said to be equal when there is no reason to believe that one rather than the other will happen as the result of any trial. But what reason could there be for such a belief based on experiment? No *a priori* principle can determine the property of a system, which is an experimental fact; we cannot tell whether a die is fair or loaded without examining it through our senses. The only experimental reason I can conceive for believing that one event is more likely to happen than another is that it has happened more frequently in the past. But if an attempt is made to define "more frequently" precisely, the judgment of equality is inevitably made to depend on the derived measurement, and the fundamental process ceases to be independent of it. This dependence is often concealed by the use of question-begging words. Thus, the principle of sufficient reason may be reasonably held to decide that, in a perfectly shuffled pack of cards, the chance that the card next after a heart is another heart is equal to the chance that it is a club. But if inquiry is made what is meant by a perfectly shuffled pack and how we are to know whether a pack is or is not perfectly shuffled, I can seen no answer except that it is one in which a club occurs after a heart as often as another heart. But, of course, to define perfect shuffling in that way is to admit that the criterion of equality is based upon the derived measurement of "frequency"*. I can find no proposed definition of the equality of chances that is both applicable to experimental facts and independent of frequency; and I conclude, therefore, that there is not for chance, as there is for resistance, a fundamental process of measurement independent of the derived.

8. But there is a further difference to be considered. Even if equality of chance could be defined independently, there would still be many chances (and those some of the most important) which could not be connected with the unit by the relations of equality and addition. Any resistance is equal to the sum of some set of resistances such that the sum of another set of them is equal to the unit or to the sum of some set of units. The analogous proposition about chances

* It is not always realized by those who calculate card chances in great detail that in actual play, even among experienced players, the shuffling is so imperfect as to distort very seriously the chances of such events as the holding of a very long suit.

is not true, if the chances are always experimentally determined. Consider, for example, the disintegration of a radioactive atom within a stated period. There are only two alternatives: the atom does distintegrate, or it does not. The sum of the chances of these two events is 1, but the chances of the two are not in general equal. And neither of them can be shown experimentally to be the sum of the equal chances of other events such that the sum of some other set of those chances is equal to the unit. The definition of unit chance together with definitions of equality and addition would never permit us to determine such chances; they can only be determined by derived measurement.

9. Chance is therefore not capable of fundamental measurement. Nevertheless the principles of fundamental measurement are important in connexion with chance, because they are involved in the calculation of chances. When we calculate a chance we always assume that it is measurable by the fundamental process. Thus if we calculate the chance of drawing a heart from a pack, we argue thus:— The chance of drawing any one of the 52 cards is equal to that of drawing any other. The chance of drawing one of the 52 cards is, by the definition of addition, the sum of the chances of drawing the individual cards, and, by the definition of unit chance, it is 1. Consequently the chance of drawing any one card is 1/52. But the chance of drawing a heart is the sum of the chances of drawing 13 individual cards; it is therefore the sum of 13 chances each equal to 13/52, *i. e.* 1/4. The calculation is perfectly legitimate, so long as we know (1) of how many individual events the event under consideration (and any other event introduced into the argument) is the sum, and (2) that the chances of these individual events are equal. (1) does not depend on the derived system of measurement, but it does involve a very complete knowledge of the event under consideration; (2), if it is an experimental proposition at all, must depend upon derived measurement. The calculation is often made when (2) is not experimental, and when there is no direct knowledge of (1); it is then purely theoretical, and the only legitimate use that can be made of it is to confirm or reject the theory by means of a comparison of the calculated chance with that determined experimentally by the derived measurement. The fact remains that true chance, the property of the system, is always and inevitably measured by the derived process and not by the fundamental.

Chance and Probability.

10. It remains to consider very briefly what connexion, if any, there is between the chance of events and the probabilities of propositions.

Probability is usually admitted to be an indefinable conception, applicable to propositions concerning which there is no complete certainty, and roughly describable as the degree of their certainty. It appears to me one of those conceptions which are the more elusive the more they are studied ; I am quite certain that I do not understand what some other writers mean by the term, and am not at all certain that I can attach a perfectly definite meaning to it myself. The observations that I can offer are therefore necessarily tentative. But it is clear, at any rate, that probability is not a property of a system and is not physically measurable : any propositions connecting it with chance must depend ultimately on fundamental judgments which can be offered for acceptance, but cannot be the subject of scientific proof.

There are two kinds of propositions the probability of which may plausibly be connected with chance ; and they naturally can apply only to systems that are characterized by chances. Of the first kind the following is typical :—This die will turn up six the next time it is thrown (or on some other single and definite occasion). Here (*cf.* P. pp. 192–200) it seems that, if the proposition is really applied to a single occasion only, the probability of the proposition must be that characteristic of absolute ignorance ; for the assumption that anything whatever is known of the result of a single trial is inconsistent with the experimental fact that the result of any one trial is random. The only exception occurs when the event is one of which the chance is so small (or so great) that the happening of it (or failure of it) would force us to revise our estimate of the chance or to deny that there was a chance at all. Of such coincidences, in systems of which the chance has been well ascertained, the assertion that they will not (or will) occur may be made with the certainty that is characteristic of any scientific statement. There is no probability.

On the other hand, it is very difficult to be sure that only a single trial is contemplated. For when such statements are important, there is always a clear possibility of a considerable number of repetitions of the trial. If this number is so great as to permit a dermination, by derived measurement, of the chance of the event within some limits relevant to the

problem, then it will be found by examination of the use of such propositions that their importance depends simply on the value of that chance. If that chance is greater than a certain value, the proposition will be true for the purposes concerned; if it is less, it will be false. I cannot myself ever find in such propositions any meaning which is not contained in the proposition :—The chance that the die will turn up six is greater or less than some other chance. Accordingly again, there seems no room for a probability which is distinct from chance.

11. Of the second kind of proposition an answer to the following question may be taken as typical :—I have two dice, of which the chances of turning up six are unequal. I throw one, but I do not know which. It turns up six. Which of the two dice have I thrown?

Here again (P. pp. 185–192), if the question is asked of a single throw, it seems to me that the only possible answer is simply, I do not know; except, as before, if the throw would be a "coincidence" with one die and not with the other. For, once more, if the events concerned are really characterized by chances, it is inconsistent with the statement that they are so characterized to assert that, at a single trial, the result, if compatible with either of the two "causes," may not happen as the result of either of them. If, on the other hand, the throws are repeated (while it is certain that the same die is always used), and if they determine the chance of one die rather than that of the other, it is clearly certain that this die, and not the other, is being used; a die can be identified by its chance as certainly as by its resistance or any other physical property. But intermediate between these extremes, there certainly seem to be cases in which, though the evidence is not sufficient to enable us to assert definitely which die is being thrown, we begin to suspect that it is one and not the other. The possibility of such a state of mind arises from the fact that there is necessarily a finite period during which the evidence is accumulating; it does not arise when, as in the usual determination of resistance, the evidence is obtained all at the same time. And our suspicion will increase generally with the "probability" as estimated by the well-known Bayes's formula for the probability of causes. In this case it appears to me that there is such a thing as probability, determined by but distinguishable from chance, and applying to a proposition, and not to an event. But I can find no reason to believe that this probability is

numerically measurable in accordance with Bayes' or any other formula.

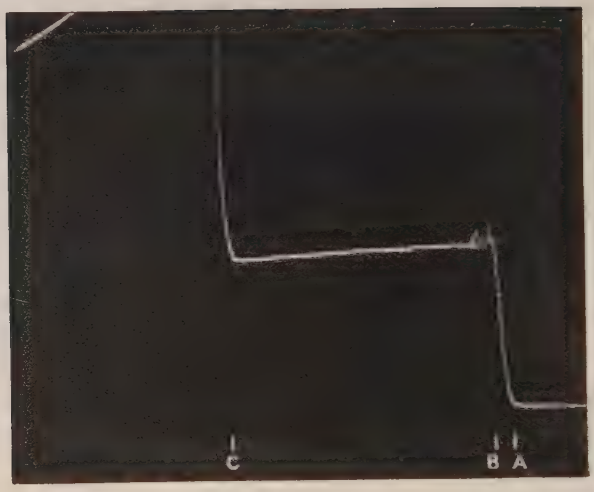
But in most cases where an attempt is made to apply a probability of causes, the condition is not fulfilled that it is known that the same die is always used. If that condition is not fulfilled, the probability, according to orthodox theory, depends on certain *a priori* probabilities which are not chances. The problem then ceases to be one of the connexion between chance and probability, and thus falls without the strict limits of our discussion.

12. But it is necessary to transgress those limits for one purpose. It has been often urged by philosophers that probability is characteristically applicable to scientific propositions, which are to be regarded, not as certain, but only as more or less probable. If this be so, the conception of chance, being a scientific conception deriving its meaning from scientific propositions, must be subsequent to the conception of probability, and the order of our discussion should have been reversed. Of course I do not accept the philosophical view, and perhaps it will be well to explain very briefly why I reject it.

Doubtless there is a sense in which scientific propositions are not certain ; but in that sense no proposition is certain, so long as its contrary is comprehensible. For if I can understand what is meant by a proposition, I can conceive myself believing it. I am not perfectly certain either that Ohm's law is true, or that $(x+a)^2=x^2+2ax+a^2$; I can conceive myself disbelieving either. If I were forced to say which I believe more certainly, I should choose Ohm's law ; for I could give a much better account of the evidence on which I believe it. A mathematician, of course, would make the opposite choice. But it appears to me useless to compare the "certainties" of two propositions when they are of so different a nature that the source of the uncertainty is perfectly different. If a proposition is as certain as any proposition of that nature can be, and if nothing whatever could make it more certain, then it seems to me misleading to distinguish its probability from certainty.

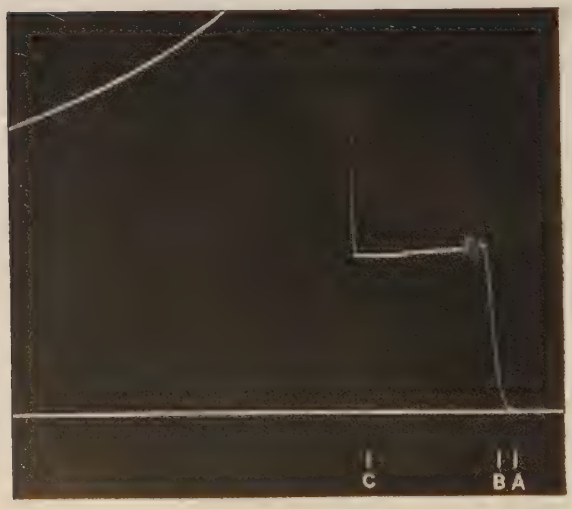
Now, fully established scientific propositions are certain in this sense. They are uncertain only in so far as they predict If in asserting Ohm's law, I mean (and I think this is my chief meaning) that it appears to me a perfectly complete and satisfying interpretation of all past experience and that other people appear to share my opinion, then Ohm's law is

Fig. 1.



Corresponding to experiment A_{17}.

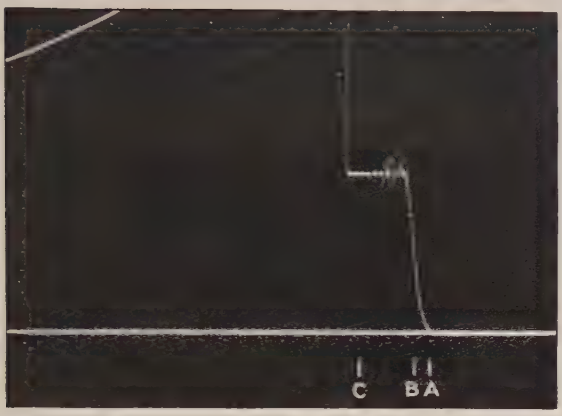
Fig. 2.



Corresponding to experiment A_{13}.

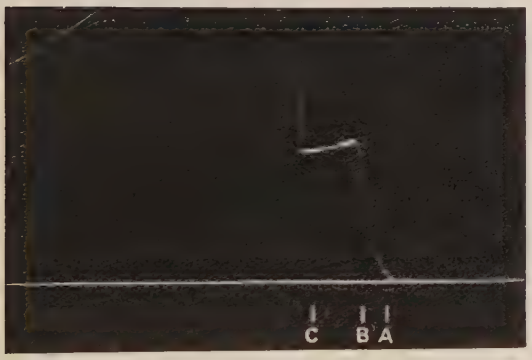
In these photographs the lower horizontal line is the line of atmospheric pressure. The ordinates represent *pressure*, and the abscissæ *time*. A is the beginning of compression, B the point of maximum compression, and C the explosion. The curve in the top left-hand corner is the cooling curve of the products of combustion.

Fig. 3.



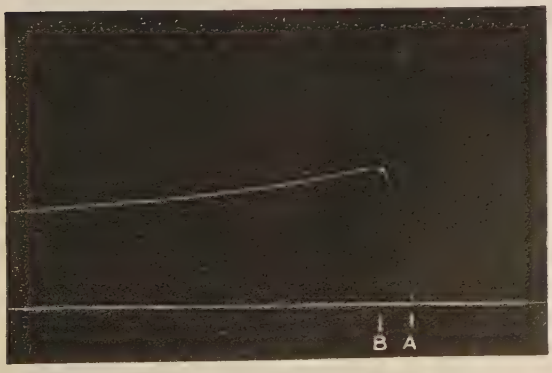
Corresponding to experiment A_{13}.

Fig. 4.



Corresponding to experiment D_3.

Fig. 5.



Corresponding to experiment D_2.

perfectly certain, or at least as certain as any mathematical or logical proposition. On the other hand, if it is meant that the law will be to me and to others an equally satisfying interpretation of all future experience, then I am not absolutely certain; I am only as certain as I can be about anything in the future. And it must be noticed that nothing can make me more certain. If I were predicting something about a single future occasion, I might in the course of time become more certain; for that future occasion might some day become past. But if, as in the case of a scientific law, I am predicting something about *all* future experience, then, since the future is indefinite, no amount of additional experience, converting finite portions of the future into the past, can make me more certain; for there will always remain as much future as before. Such uncertainty as there is in the proposition is inherent in its nature; if it were absolutely certain, it would not be the same proposition.

VIII. *Experiments on the Ignition of Gases by Sudden Compression.* *By* H. T. Tizard *and* D. R. Pye *.

[Plate I.]

1. IN a previous paper †, it was shown that when a mixture of a combustible gas or vapour with air was suddenly compressed, explosion might take place after an interval the duration of which depended on the temperature reached by the compression. It is known that below a certain temperature, called the ignition temperature, no explosion, and no very appreciable reaction, takes place under these conditions; and the experiments referred to showed that just above the ignition temperature, the delay before explosion occurs may be of the order of one second in certain cases, while—in the case of hydrocarbons and air—the delay at a temperature some 50° above the ignition temperature was very small. It was pointed out that the observed ignition temperature must not only depend on the properties of the combustible substances, but also on the conditions of experiment, and particularly on the rate of loss of heat from the gas at the

* Communicated by the Authors.

† H. T. Tizard, "The Causes of Detonation in Internal Combustion gines." Proceedings of the N.E. Coast Institution of Engineers and Shipbuilders, May 1921.

ignition temperature. The fact that this has not been fully taken into account previously seems to account in some measure for the differences in the results obtained by other workers. It was further shown that the period of slow combustion before explosion took place also depends on the properties of the combustible substance, and a theory was briefly developed connecting the "delays" observed at different temperatures with the effect of a rise in temperature on the rate of combustion, *i. e.* with the so-called temperature coefficient of the reaction. The object of the experiments described in this paper was to test these theories quantitatively, and to attempt to deduce from the results the temperature coefficients in certain typical cases. The measurement of the temperature coefficients of simple gaseous reactions is of considerable importance in connexion with the theory of chemical reactions, for one of the great difficulties in the development of theory hitherto has been the fact that most gaseous reactions have to be investigated under conditions which are complicated by the disturbing influence of solid catalysts or of the walls of the containing vessel. Gaseous reactions which occur on sudden compression are free from this complication, for the walls of the containing vessel are much lower in temperature than the gas; by quantitative measurements of the rate of loss of heat near the ignition temperature, and of the delay before explosion occurs, it therefore seems possible to gain some real insight into the mechanism of homogeneous gas reactions. Experiments of this nature also have a considerable practical interest for the development of internal combustion engines, for, according to our views, the tendency of a fuel to detonate at high temperatures depends not only on its ignition temperature but also on the temperature coefficient of its reaction with oxygen.

Previous experiments on the ignition of gases by sudden compression have been made by Falk, at Nernst's suggestion (J. Amer. Chem. Soc. xviii. p. 1517 (1906), xxix. p. 1536 (1907)), and by Dixon and his co-workers (see Dixon, Bradshaw & Campbell, Journ. Chem. Soc. 1914, p. 2027; and Dixon & Crofts, p. 2036). Nernst first put forward the view that at the ignition temperature the evolution of heat due to the reaction was just greater than that lost to the surroundings; but this suggestion has not hitherto been carried further, since no previous workers have attempted to measure the rate of loss of heat near the ignition temperature. Further, previous work has been mainly confined to the measurement of ignition temperatures of mixtures of

hydrogen, oxygen, and an indifferent gas. In such experiments the interval which occurs at the lowest ignition temperature between the end of the compression and the occurrence of ignition is very small; under these conditions an apparatus of the kind used by Nernst and Dixon gives fairly satisfactory results. It is not well suited, however, for experiments with other gases, such as the hydrocarbons, when there may be an appreciable delay before ignition occurs. In such cases it is of great importance to ensure that the cylinder in which the compression is effected is as gas-tight as possible, and that the piston is held rigidly in position at the end of the compression stroke.

II. The apparatus used for our experiments was originally designed and built by Messrs. Ricardo & Co. with a view to determining the temperatures of spontaneous ignition of various fuels used in internal combustion engines under conditions which correspond closely with those obtaining in an engine cylinder.

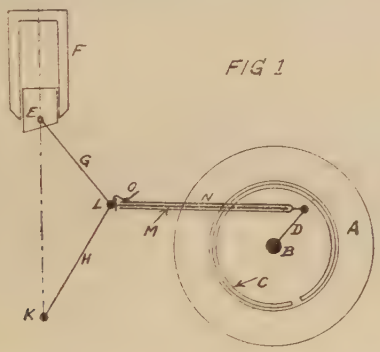


FIG 1

Fig. 1 shows diagrammatically the arrangement of the mechanism. A very heavy flywheel A rotates quite freely on the shaft B, and is kept spinning by an electric motor at about 360 R.P.M. The shaft B carries between bearings the crank D, and outside one bearing, the internal expanding clutch C, which can engage with the flywheel rim.

The piston E moves vertically in the jacketed cylinder F, which has an internal diameter of $4\frac{1}{2}$ inches and can be raised or lowered bodily in the heavy cast-iron casing of the apparatus when the compression ratio is to be altered. The length of stroke of the piston is 8 inches, and its motion is controlled by the two hinged rods G and H of which the latter is carried on a fixed bearing at K. The hinge L is

linked up with the crank pin by the compound connecting rod N. That part of the connecting rod attached to the crank pin is tubular and contains the sliding rod M attached to the hinge L. A clip O carried on the sleeve can engage with a collar on the inner rod and hold the latter rigid in the tube. With the connecting rod locked as one link, the crank is rotated by hand for setting the piston in its lowest position. When a compression has to be made, the clutch is suddenly expanded by a hand lever while the flywheel is running at high speed, clutch and crank are carried round with the flywheel, and the toggle joint ELK is straightened until the hinge L lies on the vertical line between the piston centre and hinge K. At the moment when L is vertically over K it comes up against a leather pad, and a clip comes into action which holds it in this position. At the same moment, too, the clip O releases the two parts of the compound connecting rod, so that while the two rods G and H are held in the vertical position to take the large downward thrust of the piston when explosion of the compressed mixture occurs, the crank, clutch, and flywheel are free to go on rotating, and the shock due to destruction of the momentum of the moving parts is reduced to a minimum. The initial temperature of the gases in the cylinder can be varied by means of a water jacket round the cylinder, and the variation of pressure during and after compression is recorded by means of an optical pressure indicator of the Hopkinson type.

For the purposes of the present experiments it was necessary :—

(1) To make the compression space above the piston absolutely air-tight so as to eliminate all pressure drop due to leakage.

(2) To arrange an accurate timing gear upon the indicator so that pressure-time records correct to about one per cent. could be obtained.

(3) To insert some kind of fan in the compression space so that the effect of varying turbulence in the compressed gas could be determined.

(1) *Air-tightness of the Compression Space.*

There were three points at which appreciable leakage was liable to take place :—

(*a*) Round the sides of the main piston.
(*b*) Past the indicator piston.
(*c*) Round the fan spindle.

Of these, the first two were completely eliminated and the third reduced to something quite negligible. Figures of the actual leakage are given below.

The method of eliminating piston leak will be understood from fig. 2. The piston is made in two parts. Below the

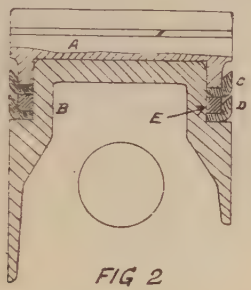


FIG 2

cast-iron top, which carries a single piston ring of the usual type, there is an aluminium body which for some distance is of smaller diameter than the cylinder. Round this waist are two cup-leathers, C and D, each with the periphery turned upwards, which are separated by a cast-iron ring of square section (E, fig. 2). The whole space above and between the cup-leathers was filled with castor oil so that no air could leak down past the piston head until it had first made a space for itself by forcing castor oil down past the two cup-leathers in series.

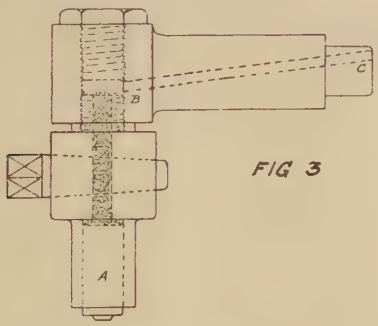
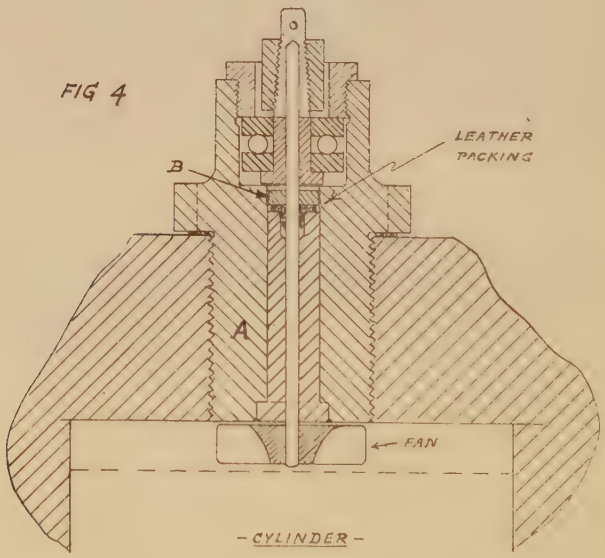


FIG 3

To eliminate leak past the indicator piston, the plan was hit upon of turning the whole indicator upside down and then pouring in a little castor oil into the space above the piston, as illustrated in fig. 3. Here A represents the piston of the inverted indicator with the space above it filled with

castor oil. The duct BC in the bracket which carries the indicator is based on the slope to prevent the possibility of any oil getting back along it and into the combustion space at C. This arrangement was very satisfactory, for, besides forming a perfect seal, it served to keep the piston well lubricated; it was found, moreover, that only a very small quantity of oil was forced down past the piston, even after continued exposure to high pressure.

The arrangement of fan and fan-spindle is shown in fig. 4. Here the plug A, shown in section, was made to screw vertically downwards in the centre of the cylinder



head, its lower end finishing flush with the top of the compression space. At the highest compression ratios this space is $\frac{1}{2}$ inch deep, so that the piston head under these conditions is close up to the under side of the fan. The fan was driven at varying speed by an electric motor through the usual form of flexible drive used for speedometers.

To prevent leakage up round the fan-spindle, this was made long and thin, and was also provided with a cup-leather of which the tightness could be adjusted by the screw B in fig. 4. By using a motor with ample reserve of power for driving the fan, it was possible to keep the cup-leather so tight as to reduce the leakage to a negligible amount.

Experiments on Air Leakage from the Cylinder.

Observations were made of the rate of fall of pressure beginning 10 seconds after compression had occurred. By this time all pressure-fall due to heat-loss has ceased, and the observed fall represents leakage only.

Experiments were made—

(*a*) With fan stationary. Compression ratio 9 : 1. Contents of cylinder were the products of combustion of a benzene and air mixture.

Time.	Deflexion on Indicator Screen.
10 secs. after compression and combustion	0·63 inches.
20 „ „ „	0·61 „
30 „ „ „	0·57 „
45 „ „ „	0·56 „
60 „ „ „	0·53 „

The indicator calibration was 1 inch deflexion = 188 lb. per sq. in., so that the above gives a rate of fall of $\frac{0{\cdot}1 \times 188}{50} = {\cdot}38$ lb. per sq. in. per sec.

(*b*) With fan rotating at full speed, 2000 r.p.m. Compression ratio 10 : 1. Contents of cylinder pure hydrogen.

Time.	Deflexion on Indicator Screen.
10 secs. after compression	0·66 inches.
20 „ „	0·53 „
30 „ „	0·40 „

This, owing to the low viscosity of hydrogen and the high compression ratio, was a very severe test. The rate of fall was $\frac{{\cdot}26 \times 188}{20} = 2{\cdot}45$ lb. per sq. in. per sec. Now in this experiment the rate of fall of pressure due to *cooling* was as high as 600 lb. per sq. inch per second immediately after compression, so that the effect of gas leakage on the apparent rate of cooling was clearly negligible.

The Indicator.

It will be convenient at this point, before going on to describe the timing apparatus, to give some further data as to the indicator used. As stated above, this was of the

standard Hopkinson optical type *, and need not be described in detail. The piston used throughout the experiments was one of 0·125 sq. in. area. Pressure on the piston deflects a spring attached to it, and thus tilts a small mirror which reflects a point of light from a fixed lamp. The magnitude of the pressure is thus arranged to correspond to the vertical (downward) deflexion of the image of the light on a $4\frac{1}{4} \times 3\frac{1}{4}$ photographic plate carried in a camera fixed to the indicator.

The calibration of the pressure scale was made in two ways. Firstly, by subjecting the piston to known oil pressures and measuring deflexions on the camera screen, and, secondly, as a check on the piston area, by direct dead-weight loading of the spring and measurement of the deflexion. The oil pressures in the first calibration were produced by a carefully gauged vertical plunger loaded with weights. The calibration by this method was carried out at the beginning of the experiments, and it was found that the relation between pressure and deflexion was given very closely by a straight line at a slope corresponding to a pressure rise of 188 lb. per sq. in. per inch deflexion. At the conclusion of the experiments this calibration was checked by the dead-weight calibration as follows:—

Load on Spring.	Reading on Camera Screen. Loading.	Unloading.
0	0·39 in.	0·40 in.
Weight carrier = 3·25 lb.	0·54 „	0·54 „
Carrier + 10 = 13·25 „	0·975 „	0·99 „
„ + 20 = 23·25 „	1·40 „	1·43 „
„ + 30 = 33·25 „	1·81 „	1·81 „

This also shows a satisfactorily straight-line relationship which checks remarkably well with the previous calibration, for

$$\text{Dead weight per inch deflexion} = \frac{33{\cdot}2}{1{\cdot}42} = 23{\cdot}4 \text{ lb.}$$

corresponding to a pressure on a piston of area 0·125 sq. in. of $8 \times 23{\cdot}4 = 187{\cdot}2$ lb.

The previous figure of 188 lb. per sq. in. per inch deflexion has been used throughout the calculations.

(2) *The Timing Gear.*

The pressure in the cylinder being given by the vertical motion of a point of light on a photographic plate, it was necessary for the accurate measurement of the lengths of the delay period to give the point of light at the same time

* Hopkinson, Proc. Inst. Mech. Eng. Oct. 1907, p. 863.

a uniform and exactly known motion across the plate horizontally. To obtain this uniform horizontal velocity, the arrangement in fig. 5 was adopted.

A is a vertical cylinder closed at the bottom and carrying at the top a guide for the piston rod B. At the lower end of the rod B is a light loosely fitting piston in which are drilled one or more small holes. The rod B has a collar C above the guide, and can be loaded with weights, up to 50 lb. if need be. The cylinder is filled with paraffin which, when the piston and weights have been raised and then released by a trigger arrangement, is forced at high velocity through the holes in the piston from the under to the upper side. This arrangement gives a velocity of fall of

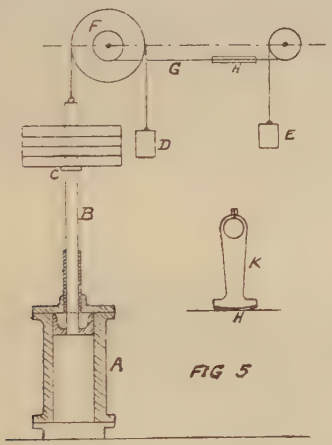


FIG 5

the piston which is uniform to within 1 or 2 per cent., and it was found, moreover, that this velocity varied directly as the square root of the weight carried : a result which shows that the viscosity of the oil is a negligible factor in determining the rate, and that the latter will therefore not be affected by any small changes of temperature which might occur from day to day in the oil.

This point was checked experimentally, and it was found that the maximum change in the rate of fall produced by heating the oil from 19° C. to 65° C. was 1½ per cent. As the oil temperature during the experiments never varied by more than 2° or 3°, the effects of temperature changes were quite negligible.

Piston and weights are suspended by a steel wire from an arrangement of pulleys (F) carried on spindles attached to

the cylinder head. D and E are balance weights to keep the wires taut, and F is a compound pulley which reduces the horizontal motion of the wire G to about one-half the vertical movement of the piston. At H the wire G is divided, and each half wraps round the periphery of a sector attached to the arm K of the indicator. The barrel of the indicator is thus uniformly rotated through a sector about a vertical axis, and the speed of the point of light horizontally across the photographic plate in the indicator camera can be varied between wide limits by alteration of the weights and the number of holes in the loaded piston through which the paraffin is forced. Actually it was found in these experiments that only two speeds were required, of 5·77 and 4·30 cm. per second. These were obtained when the piston carried weights of 40 and 24 lb. respectively, and was pierced by a single hole $\frac{1}{4}$ inch diameter.

Both the speed and its uniformity were measured by interrupting the light at known intervals while it traced a straight horizontal line across the plate. To obtain the uniform interruptions, 60 equally spaced contacts were arranged round the periphery of a disk about 6 inches in diameter. An electric motor was used to rotate the disk through a friction clutch, the speed being kept steady by the operation of a governor, which caused the disengagement of the clutch when the speed tended to increase. The arrangement was very similar to that of a gramophone motor, except that the speed of the latter is kept steady by means of a little brake which is operated by a governor, whereas in our experiments the governor operated the clutch. This apparatus, which was made at the Royal Aircraft Establishment, Farnborough, was lent to us for these experiments by kind permission of the Director of Research, Air Ministry.

To calibrate the falling weight apparatus, the speed of revolution of the disk was first adjusted roughly to one complete revolution in 2 seconds. It was then timed repeatedly over twenty complete revolutions by stop-watch; the mean time for this number of revolutions was found to be 36·5 seconds, the variations in successive timings not exceeding ·4 second. Hence, since sixty contacts were made each revolution, the time between the beginnings of successive contacts was $\frac{36{\cdot}5}{60 \times 20} = 0{\cdot}0304$ second. The 6-volt lamp of the indicator was then connected up, through the contacts on the disk, to a 12-volt battery; it was thus greatly "overrun" whenever contact was made. The effect of

overrunning the lamp in this way was to make it flicker brightly, so that when the indicator mirror was rotated by the dropping weight, the reflected light made a horizontal line of dashes across the photographic plate, the interval between the middle of two consecutive dashes corresponding to the interval between two contacts on the uniformly rotating disk. The magnitude and uniformity of the velocity of rotation of the indicator mirror by the falling weight could thus be measured. Records were taken with falling weights of 8, 24, and 40 lb.

To show the uniformity of motion during descent, the results of measuring two plates taken with the 24-lb. load are given in full. The apparent variations in any single experiment are due rather to the difficulty of estimating the centre of a flicker to 0·1 mm. than to any real variation in the rate of fall of the loaded piston.

	Plate A.	Plate B.
5 flickers in	0·70 cm.	0·65 cm.
10	1·34	1·30
15	2·0	1·95
20	2·66	2·6
25	3·30	3·28
30	3·98	3·95
35	4·63	4·58
40	5·29	5·22
45	5·93	5 85
50	6·59	6·51
55	7·22	7·12
60	7·90	7·78
65	8·53	8·40
70	9·21	9·06
75	9·89	9·70
Mean 1 flicker in	0·132 cm.	Mean 1 in 0·129 cm.
1 flicker =	·0304 second.	
∴ 1 second =	4·34 cm. on plate.	1 second = 4·26 cm.

Mean 4·30 cm. = 1 second.

Similar records with loads of 8 and 40 lb. gave time scales of

2·22 cm. = 1 second

and 5·77 cm. = 1 second.

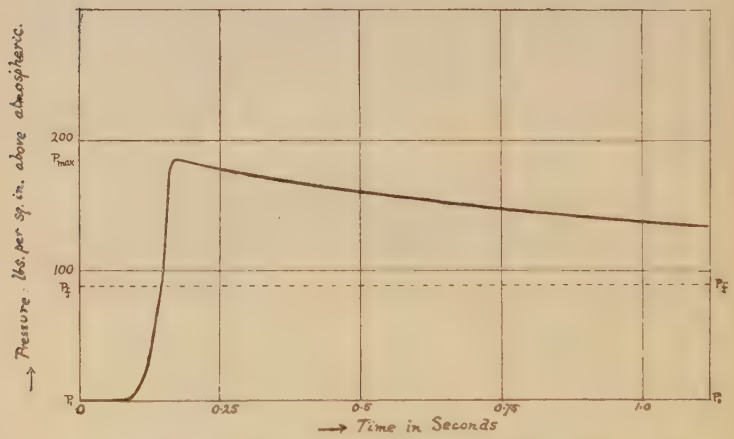
The means for the two plates with 40-lb. load were 5·75 and 5·79 cm. per second, so that it may safely be assumed that the time scale was known correctly to within 2 per cent., and probably less. This is a degree of accuracy as great as that with which, as a rule, it is possible to measure the records.

The vertical movement of the weights was about 7 inches,

which produced a horizontal movement of the point of light more than twice as great as was necessary to traverse the photographic plate, so that ample time was provided for the speed to have become uniform before a pressure record was taken. The fact that the indicator and all gear except the actual falling weight was carried on the cylinder head, made it possible to raise or lower this when changing the compression ratio without affecting the timing gear. The only difference was an alteration of about an inch in the distance fallen by the weights between the position for maximum and minimum compressions.

Figs. 6 and 7 are prints taken from two typical records.

Fig. 6.

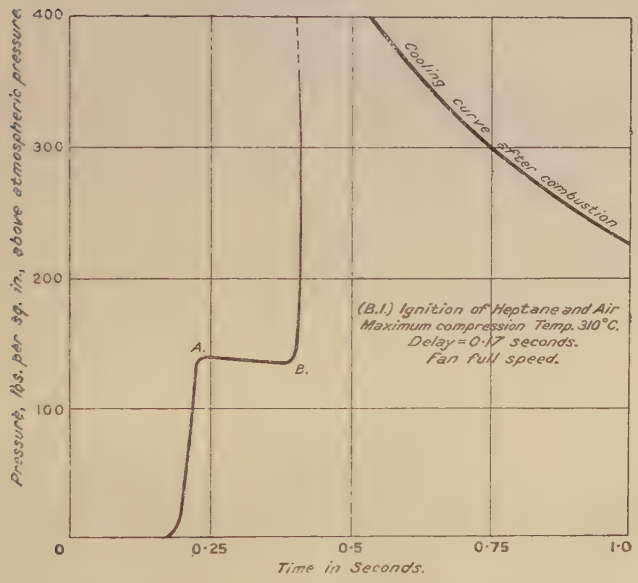


In the first there is no explosion, and the record shows simply the rise of pressure due to compression, and subsequent fall as the compressed mixture cools. Fig. 7 shows the first rise of pressure due to compression, the "delay" period of nearly constant pressure followed by the practically instantaneous rise on explosion, and finally the rapid cooling of the intensely hot products of combustion. It may be mentioned that the spring of the indicator was protected from the force of the explosion pressure by a stop which prevented the deflexion from ever being greater than just to the edge of the photographic plate.

III. If a gas at a known temperature and pressure is suddenly compressed in a gas-tight cylinder, we can calculate from a measurement of the maximum pressure reached the average temperature of the gas at the moment of maximum

compression. If the volume then remains constant, measurement of the rate of fall of pressure with time gives the rate of loss of heat, if the specific heat of the gas is known. One would expect this rate to be closely proportional at any moment to the difference in the average temperature of the

Fig. 7.



gas and that of the cylinder walls. If the temperature of the gas is T° absolute, and that of the walls is θ, then

$$-\frac{dT}{dt}=\alpha(T-\theta), \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where "α" is a constant which we call the cooling factor, and which depends on the nature of the gas, and its degree of turbulence in the cylinder. The results given later will be seen to justify this equation.

Integration of (1) gives

$$\log_e \frac{T_0-\theta}{T-\theta}=\alpha t, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where T is the average temperature when t seconds has elapsed from the moment the average temperature was T_0.

Since the total volume remains constant, this equation can be written

$$\log_e \frac{P_{max.}-P_f}{P-P_f} = at \quad . \quad . \quad . \quad . \quad . \quad (3)$$

if the simple gas laws hold. $P_{max.}$ is here the maximum compression pressure, P the observed pressure t seconds after the maximum compression pressure is reached, and P_f the final pressure of the gas when its temperature is the same as that of the walls. P_f is therefore equal to $r \times P_1$, where r is the compression ratio and P_1 the initial pressure (in these experiments one atmosphere) before compression.

Fig. 6 is a typical cooling curve obtained when air at an initial pressure of 14·73 lb. sq. in. and temperature of 23° C. was compressed in the ratio of 7·02 to 1. The time of compression was 0·08 second, and the values of the cooling factor a, obtained from observations of the fall in pressure, are given in the following table :—

TABLE I.

Time after max. pressure (seconds).	Observed pressure = P lb./sq. in.	$P-P_f$.	$\log \frac{P_{max.}-P_f}{P-P_f}=\delta$.	$a=\frac{2\cdot3\delta}{t}$.
0	199·7	96·3		
·173	185·7	82·3	·068	·90
·347	174·6	71·2	·131	·87
·52	165	61·6	·194	·86
·695	158·3	54·9	·244	·81
·87	152·4	49	·294	·78
				Mean=·85
P_1=14·73 lb./sq. in.	θ (initial temp.) = 23° C. Max. temp. (calc.) = 299° C.		$P_f=P_1\times7\cdot02=103\cdot4$.	

It was found as a rule that the calculated cooling factor tends to diminish as "t" increases. This may be due in part to errors of observation, for the errors are big when the time interval after the attainment of maximum pressure is small ; but experimental error will not wholly account for it, and it may be explained on the reasonable assumption that there is a fair degree of turbulence in the gas just after a sudden compression, which dies down after a short time. Rotation of the fan at a high speed increases the rate of cooling considerably ; experiments with air gave a cooling factor about

three times as great when the fan was running as when the gas was "stagnant." This will be referred to later when discussing the results of the experiments of ignition. Raising the compression ratio also increases the cooling factor for air. This is also to be expected since the distance between the top of the piston and the head of the cylinder is lessened. The cooling factor for "stagnant" air in the apparatus was found to increase from about 0·76 at a compression ratio of 6 to 1 to about 1·0 at a compression ratio of 10 to 1, the distance between the piston and cylinder head being approximately 4 cm. (1·6 in.) at the lower and 2·3 cm. (·9 in.) at the higher compression ratio. The experimental error in the cooling factors, obtained by experiments similar to that quoted above, is probably about 5 per cent. It should be noted that the rate of loss of heat under the above conditions is considerable. For instance, in the experiment quoted the maximum difference in temperature between the gas and the walls after compression is 276°; and since the specific heat c_v is about 0·18, the air is losing heat at the maximum temperature at the rate of

$$0{\cdot}18 \times {\cdot}85 \times 276 = 42 \text{ calories per second per gram.}$$

It follows that if an explosive mixture of gases is suddenly compressed to its ignition temperature, in such an apparatus as that described above, the initial rate of the chemical reaction at the lowest temperature at which ignition is observed must be considerable, for the evolution of heat due to the reaction must equal approximately the rate at which heat is lost to the walls. For example, the total heat of combustion of a mixture of a paraffin hydrocarbon and air, in the correct proportions to burn to CO_2 and water, is approximately 700 calories per gram of mixture. If it ignites when suddenly compressed to such a temperature that the rate of loss of heat is 35 calories per gram, the reaction, if it continued uniformly at the initial rate, would be complete in 20 seconds. This illustration may serve to show the general nature of the reactions that occur on sudden compression; what occurs in practice is that the gas, or part of the gas, reacts so that the evolution of heat takes place at a somewhat higher rate than the loss of heat by conduction etc.; hence the temperature of the reacting gases must automatically increase, and with it the rate of reaction, until the gas "explodes." The interval between the end of compression and the explosion must clearly depend mainly on three factors: (*a*) the compression temperature, (*b*) the temperature coefficient of the reaction, (*c*) the rate of loss of heat to the walls.

IV. It is now generally recognized that the rate of an ordinary chemical reaction varies with the temperature in a way which may be empirically expressed by an equation of the form

$$k = Ae^{-\frac{B}{T}}, \quad \ldots\ldots\ldots \quad (4)$$

where k is the velocity constant, and A and B are constants, B being the temperature coefficient.

In the case of reactions evolving heat, we can write this:

$$Q = Ae^{-\frac{B}{T}}, \quad \ldots\ldots\ldots \quad (5)$$

where Q is the initial rate of evolution of heat when a definite mixture is suddenly compressed to the temperature T. At the lowest ignition temperature T_0 we have,

$$Q_0 = Ae^{-\frac{B}{T_0}}, \quad \ldots\ldots\ldots \quad (6)$$

where Q_0 can be measured from observation of the rate of cooling of the gases at a temperature slightly below the ignition point: *i. e.*,

$$Q_0 = \alpha c_v (T_0 - \theta), \quad \ldots\ldots\ldots \quad (7)$$

where α is the cooling factor, c_v the specific heat of the mixture, and θ the temperature of the walls.

Now suppose that the gases are compressed initially to a higher temperature T: the initial rate of loss of heat will then be higher than Q_0, namely

$$Q_0 \frac{T - \theta}{T_0 - \theta};$$

but since the effect of temperature on the rate of evolution of heat due to the reaction is so much greater than that on the rate of loss of heat, it is sufficiently accurate at present to assume for our purposes a constant initial rate of loss of heat $= Q_0$. The initial rate of reaction at the higher temperature will be given by

$$Q_T = Ae^{-\frac{B}{T}}.$$

From (5) and (6) we have

$$\log \frac{Q_T}{Q_0} = \frac{b}{T_0}\left(1 - \frac{T_0}{T}\right), \quad \ldots\ldots \quad (8)$$

where $b = 0{\cdot}4343\,B$.

Now under these circumstances the net rate of gain of heat is $(Q_T - Q_0)$ calories per second.

The initial rise of temperature is therefore given by

$$c_v \frac{dT}{dt} = Q_T - Q_0 = Q_0\left(\frac{Q_T}{Q_0} - 1\right), \quad . \quad . \quad . \quad (9)$$

or
$$\frac{1}{\left(\frac{Q_T}{Q_0} - 1\right)} . dT = \frac{Q_0}{c_v} dt,$$

or
$$\frac{1}{\left(\frac{Q_T}{Q_0} - 1\right)} . d\left(\frac{T}{T_0}\right) = \frac{Q_0}{c_v T_0} dt.$$

But since
$$Q_0 = \alpha c_v (T_0 - \theta),$$

we have

$$\frac{1}{\left(\frac{Q_T}{Q_0} - 1\right)} . d\left(\frac{T}{T_0}\right) = \alpha\left(\frac{T_0 - \theta}{T_0}\right) dt. \quad . \quad . \quad (10)$$

V. Can this equation be used to determine the delay, or period of slow combustion, that occurs before the temperature suddenly rises very rapidly, *i.e.* before explosion takes place? Strictly speaking, this could only be done if Q, the rate of evolution of heat during the initial reaction, depended only on temperature and not on the concentration of the reacting substances. According to all ordinary theories of reaction, this would not be true; the rate of reaction should depend in some way on the concentration of the reactants. If we consider such a reaction as the combustion of heptane

$$C_7H_{16} + 11\,O_2 = 7\,CO_2 + 8\,H_2O,$$

we can hardly suppose it is necessary, before the initial reaction—whatever is its nature—can occur, for 1 molecule of heptane and 11 molecules of oxygen to collide; but it is reasonable to assume that the rate of reaction must at least be proportional either to the concentration of the heptane or to that of the oxygen, or to the product of the two. It is necessary, however, to point out that there is little, if any, satisfactory evidence that homogeneous reactions in gases obey what is ordinarily understood by the law of mass action. In fact, evidence from the ignition of gaseous mixtures by sudden compression points rather to the reverse. Dixon and Crofts' experiments * on the ignition of mixtures of hydrogen and oxygen are difficult to explain by any reasonable assumption as to the mechanism of the reaction based on the

* Dixon & Crofts, Trans. Chem. Soc. 1914, p. 2036.

law of mass action, even after taking into account the possible effects of a different rate of loss of heat in the mixtures with which they experimented ; the authors state, in fact, that they "can offer no satisfactory explanation of the phenomena observed." Recent experiments by one of us * have shown that the ignition temperatures of hydrocarbon-air mixtures are independent, within the errors of measurement, of the proportion of the hydrocarbon within quite wide limits. For instance, when the proportion of heptane to air was varied in the ratio of 1 to 10, the ignition temperature was only lowered apparently by some 8° C. from 293° C. to 285° C. As the error of observation may certainly amount to 4° C. in such experiments, there is here no evidence that the alteration of the concentration of heptane has any effect on the ignition temperature. If, as might appear reasonable on the law of mass action, the rate of reaction depended directly on the concentration of the heptane, we should expect it to be 10 times as great in one case as in the other. Now, it is shown later in this paper that the temperature coefficient B, in the case of heptane, is about 13,000, from which it follows from equation (5) or (8) that if the initial rate of reaction on compression were really 10 times as great in one case as in the other, the ignition temperature of the richer mixture should be over 50° lower than that of the weaker mixture. One could not fail to detect with certainty a difference of this order : hence from the experimental results we come to the conclusion that the rate of combustion under these conditions does not depend on the concentration of the heptane vapour within wide limits. Experiments with other similar substances support this view.

VI. If this be correct, the only probable alternative is that the rate depends essentially on the concentration of the oxygen. We have not yet attempted to put this to a direct test in the apparatus used by us, since it would be necessary, to do this completely, to explode detonating mixtures of hydrocarbons and oxygen ; and we were anxious to avoid the danger of breaking the apparatus before other important experiments were carried out. We intend, however, to examine this point in the near future.

It remains to consider whether, if the rate of reaction is directly proportional to the concentration of oxygen, the effect of the automatic decrease in oxygen content on ignition can be safely left out of account in calculating "delays" by

* Tizard (*loc. cit.*).

means of equation (10). For this purpose we have to estimate to what extent the reaction occurs during the slow period of combustion before the explosion occurs, and the pressure-time curve becomes almost perpendicular to the time axis (see fig. 7). Taking the experiments on heptane as an example, we find (Table III.) that the biggest delay observed was about 0·6 second; that the ignition temperature under these conditions is 280° C. = 553 absolute, and that the difference of temperature between the gas and the walls was about 240° C. The observed cooling factor was 0·51. Since the specific heat of the mixture experimented with is approximately = 0·2, the rate of reaction at the ignition temperature must correspond to a heat evolution of

$$0{\cdot}51 \times 0{\cdot}2 \times 240 = 24 \text{ calories per gram per second.}$$

The total heat of combustion of 1 gram of the mixture (containing about 75 per cent. of the theoretical amount of heptane for complete combustion) is about 510 calories, so that it is evident that during the period when the temperature is only rising slowly, which is always less than half a second, the amount of reaction and therefore the changes in concentration must be small. Once the temperature begins to rise quickly, it is evident that the disappearance of oxygen can have only a secondary effect on the rate of the reaction compared with that of the rise of temperature until the combustion is nearly complete, so that the error involved in the calculation from equation (10) of the total time of combustion, by ignoring the effect of changing concentration, must be small. In view of the unavoidable experimental errors in carrying out experiments of this kind and of our still incomplete knowledge of the mechanism of combustion, we do not, in fact, think that any attempt to take fully into account such secondary effects is justified at present.

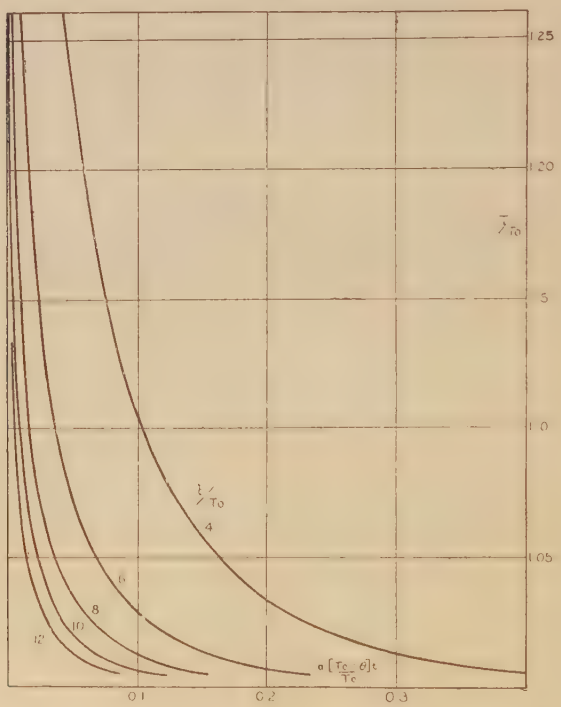
VII. We therefore arrive at the conclusion that the time for complete combustion of an explosive mixture of gases when suddenly compressed to a temperature above its ignition temperature is closely given by the integration of equation (10), and is therefore

$$t = -\frac{1}{\alpha\left(\frac{T_0 - \theta}{T_0}\right)} \int_{T}^{T_0} \frac{1}{\frac{Q_T}{Q_0} - 1} d\left(\frac{T}{T_0}\right), \quad . \quad . \quad (11)$$

where T_0 is the lowest ignition temperature under the conditions of the experiment, and T is the initial compression temperature.

In this equation, $\frac{Q_T}{Q_0}$ is given by equation (8); its value depends on the magnitude of the temperature coefficient of the reaction. It is not possible to integrate the equation completely, but the integration can be carried out by approximate methods for any value of $\frac{b}{T_0}$. In the table below are given values of the expression $\alpha\left(\frac{T_0-\theta}{T_0}\right)t$ at different values of $\frac{T}{T_0}$, and for various values of $\frac{b}{T_0}$. The corresponding curves are shown in fig. 8.

Fig. 8.



If our views are correct, these should form standard curves representing the delays which should be observed under different conditions when any explosive mixture of gases is compressed to a temperature above its lowest ignition temperature. The application of the theory to any specific case should enable the temperature coefficient of the reaction to be determined.

TABLE II.

Values of $\alpha\left(\frac{T_0-\theta}{T_0}\right)t$ for different values of $\frac{T}{T_0}$ and $\frac{b}{T_0}$.

T/T_0.	b/T_0=12.	=10.	=8.	=6.	=4.
1·004	·086	·112	·155	·233	·418
1·01	·056	·075	·108	·169	·321
1·03	·0245	·036	·056	·098	·209
1·07	·0081	·014	·025	·051	·131
1·11	·0033	·0065	·0135	·032	·095
1·27	·0002	·0007	·0022	·0083	·0415

The curves bring out clearly the effect of the two main factors which determine the characteristics of an explosion by sudden compression; namely, the initial rate of loss of heat, and the temperature coefficient. If two gases have the same ignition temperature under the same conditions of loss of heat, the sharpest explosion will occur in the case of the gas with the highest temperature coefficient, and the greater in this case will be the effect, on the magnitude of the delay before explosion, of a higher temperature of compression. On the other hand, in any one case, the ignition temperature will be raised by carrying out the experiment under conditions which involve an increased rate of loss of heat; at the same time the sharpness of the explosion will also be increased.

VIII. To test the above theory, and to use it to obtain a measure of the temperature coefficient in certain cases, we chose three substances: heptane C_7H_{16}, ether $C_2H_5.O.C_2H_5$, and carbon bisulphide CS_2. These substances were chosen for the following reasons: (*a*) they could be obtained in a sufficiently pure state; (*b*) they all have low ignition temperatures, which lessens the practical difficulties of the experiments; (*c*) they are known to behave very differently from the point of view of detonation when used as fuels for internal combustion engines; and (*d*) their difference in chemical and physical properties makes the comparison of their behaviour on combustion particularly interesting. To test the theory adequately we considered it absolutely necessary, particularly in view of the simplifying assumptions made, not to be content with one set of conditions for the ignition experiments. Two series of experiments were therefore made with each substance; in the first series the

gaseous mixtures with air were compressed in a non-turbulent condition, while in the second a high-speed fan was kept running throughout the period of compression and subsequent ignition. The use of the fan increased the rate of heat-loss at the compression temperature by about three times; hence the difference in ignition temperatures observed with and without the fan running gives an important and necessary check on the value of the temperature coefficient which is calculated from the "delay" curve obtained when the non-turbulent gases are compressed. The temperatures given below represent the average temperature of the gas at the instant of maximum compression. By measuring the compression pressures in each experiment, a value of "γ," the apparent ratio of the specific heats, is obtained from the expression:

$$\frac{P_{max.}}{P_1} = r^{\gamma},$$

where r is the compression ratio.

The average value of γ is taken for the series, and the compression temperatures then calculated for each case from the expression

$$\frac{T_{max.}}{\theta} = r^{\gamma - 1}.$$

In each set of experiments the initial mixture of gases was of the same composition throughout, the proportion of air being somewhat greater than that required for complete combustion. The intention of using a weak mixture was to avoid as far as possible the deposition of carbon; as stated above, the absolute ignition temperature is not affected appreciably by fairly wide changes in the original strength of the mixture.

IX. The first results with heptane gave a very satisfactory confirmation of the theory developed above. The results of the experiments are given in the following tables and diagrams, which include measurements from *all* the records made under each set of conditions. No unsatisfactory records have been discarded.

TABLE III.

Ignition of mixtures of heptane and air.

Mean apparent value of "γ" observed = 1·313.

Fan stationary. Initial pressure (atmospheric) = 14·8 lb./sq. in. in expts. A_1 to A_{10}, and 14·9 lb./sq. in. in expts. A_{11} to A_{17}.

Strength of mixture = 1 grm. heptane : 20 grms. air.

No. of expt.	Compn. ratio.	Initial temp.	Max. Compn. pressure.	Max. Avge. temp. calc. from $\gamma=1{\cdot}313$.	Delay obs.	Cooling factor $=a$.
1	5·55	51° C.	140·5 lb./sq. in.	280° C.	No ignition.	0·49
2	6·03	50·5	156	295	0·19 sec.	—
3	7·02	50·5	192	323	0·04	—
4	8·02	49·5	225	346	0·007	—
5	5·55	46	141	273	No ignition.	0·53
6	5·82	46	151·5	281	0·56	—
7	6·23	46	163	293	0·21	—
8	8·02	46·5	227	340	Very small.	—
9	7·02	47	189·5	316	0·06	—
10	7·02	42	166	307	0·07	—
11	6·56	58	174	324	0·05	—
12	6·56	53	177	315	0·06	—
13	6·56	48	176	306	0·12	—
14	6·56	44	177	298	0·18	—
15	6·56	41	176	293	0·28	—
16	6·56	38	178	288	0·25	—
17	6·56	35	176	282	0·58	—

The cooling factors in Table III. were obtained from the results of those experiments where no ignition occurred by the application of equation (3). The results for A_1 were as follows :—

TABLE IV.

Calculation of cooling factor.

Time from max. pressure.	Obs. pressure. lb./sq. in.	$P-P_f$.	$\log\frac{P_{max.}-P_f}{P-P_f}=\delta$.	$a=2{\cdot}3\frac{\delta}{t}$.
0	140·5	58·5	—	—
0·23 sec.	134	52	·0512	0·51
0·47	127·5	45·5	·109	0·53
0·7	123·5	41·5	·149	0·49
0·93	119	37	·199	0·49
1·16	116	34	·236	0·47
1·39	112·5	30·5	·283	0·47

Mean = 0·49

$$P_f = 5{\cdot}55 \times 14{\cdot}8 = 82{\cdot}1 \text{ lb./sq. in.}$$

The results for plate A_5 were similar, the mean being $\alpha = 0{\cdot}53$. In the calculations the figure 0·51 has been taken.

TABLE V.

Ignition of mixtures of heptane and air. Fan full speed. Initial pressure = 14·86 lb./sq. in. Strength of mixture as before. Mean apparent value of $\gamma = 1{\cdot}310$.

No. of expt.	Compn. ratio.	Initial temp.	Max. Compn. pressure.	Max. Aver. temp. calc.	Delay obs.	Cooling factors.
B_1 ...	6·03	61° C.	157·5	310° C.	0·16 sec.	—
B_2 ...	5·52	59	139	291	No ignition.	1·43
B_3 ...	5·73	59·5	148	298	,,	1·40
B_4 ...	5·91	59·5	154	304	,,	1·36
B_5 ...	6·13	59·5	161	310	,,	1·47
B_6 ...	6·33	59·5	165	313	0·13	—
B_7 ...	6·56	59	175	321	0·08	—
B_8 ...	7·01	60	188	336	0·05	—
B_9 ...	8·02	60	?	362	0·00*	—

Mean = 1·42.

* Ignited before top of compression.

X. Considering firstly the results of the experiments without the fan running, we find that at the lowest ignition temperature of about 280° C., the temperature of the walls was about 40° C. The expression $\alpha \frac{(T_0 - \theta)}{T_0}$ (equation 11) has therefore the value—

$$\frac{0{\cdot}51 \times 240}{553} = 0{\cdot}22.$$

Assume that the true ignition temperature under these conditions is 280° C. = 553 absolute, and that $\frac{b}{T_0} = 10$: then by integration of the theoretical equation (see Table II.) we obtain the following results :—

TABLE VI.

Ignition of heptane by air. No fan.

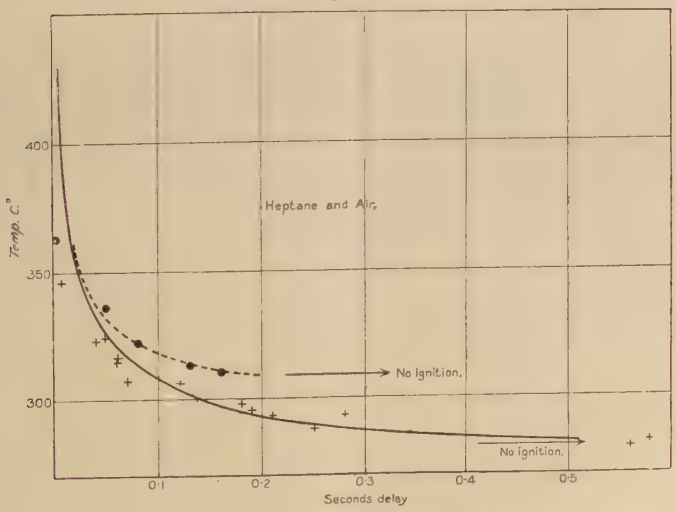
Lowest ignition temperature T_0 taken as 553 absolute.

$\frac{b}{T_0}$ assumed = 10.

T/T_0.	T.	$t°$ C.	$a\left(\frac{(T_0-\theta)}{T_0}\right)t$ (theor.).	t (calc.).
1·004	555	282	0·112	0·51 sec.
1·01	558·5	285·5	0·075	0·34 ,,
1·03	569·5	296·5	0·036	0·164 ,,
1·07	592	319	0·014	0·063 ,,
1·11	614	341	0·0065	0·03 ,,
1·27	702	429	0·0007	0·003 ,,

The last column contains the theoretical "delays" that should occur at compression temperatures given in the third column, if the temperature coefficient of the reaction corresponds to a value of $\frac{b}{T_0} = 10$.

Fig. 9.



The corresponding theoretical curve is shown in fig. 9, the experimental points taken from Table III. being marked with a cross. The general agreement is all that could be desired.

XI. If the value for the temperature coefficient so deduced is correct, it should be possible to use it to calculate the higher ignition temperature when the fan is used, and also the shape of the new delay curve.

Now, the mean cooling factor with the fan has been shown to be 1·42. The ignition temperature under these conditions is evidently about 310° C. (see Table V.), the temperature of the walls being 60° C. Hence the ratio of the rate of loss of heat with and without fan at the respective ignition temperatures is—

$$\frac{Q_0'}{Q_0} = \frac{1{\cdot}42\,(310-60)}{0{\cdot}51\,(280-40)} = 2{\cdot}90.$$

Hence the new (theoretical) ignition temperature T_0' (with fan) should be given by (see equation (8))—

$$\log \frac{Q_0'}{Q_0} = \log 2{\cdot}90 = \frac{b}{T_0}\left(1 - \frac{T_0}{T_0'}\right):$$

$$i.\,e.\quad 0{\cdot}462 = 10\left(1 - \frac{T_0}{T_0'}\right)$$

$$\text{or}\qquad \frac{T_0}{T_0'} = 0{\cdot}954.$$

$$T_0' = 580 = 307°\ \text{C}.$$

This is close to the observed value.

The new value of $\frac{b}{T_0}$ will be $\frac{5530}{580} = 9{\cdot}5$: while the value of $\alpha\left(\frac{T_0' - \theta}{T_0'}\right) = \frac{1{\cdot}42 \times 247}{580} = 0{\cdot}60.$

The corresponding theoretical values of the delay are shown in the following table.

TABLE VII.

Ignition of heptane by air. Fan full speed.

Ignition temperature calculated from previous results, 580°C.

$$\frac{b}{T_0'} = 9{\cdot}5.$$

T/T_0.	T.	t° C.	$\alpha\left(\frac{T_0' - \theta}{T_0'}\right) t$ (theor.).	t (calc.).
1·004	582	309	0·12	0·20 sec.
1·01	586	313	0·08	0·13
1·03	597	324	0·04	0·07
1·07	621	348	0·016	0·027
1·11	644	371	0·0076	0·013

The theoretical values given in this table are represented by the dotted curve in fig. 9, the experimental results (Table V.) being shown by circles. The close agreement between experiment and theory is obvious; it is, indeed, closer than could reasonably have been expected in view of the fact that the temperature errors must be estimated as 3 or 4 degrees, while the cooling factors are subject to an error of about 5 per cent. The results can, however, leave little doubt of the substantial accuracy of the simple theory worked out above, and the temperature coefficient deduced must be very near the truth. It is of great interest to note that it is of the same order as that of chemical reactions in liquids at ordinary temperature; for the reaction velocity is approximately doubled for a 3 per cent. rise in absolute temperature.

XII. The only experimental values given in Tables III. and V. which seem to call for any special comment are those corresponding to experiments A_4 and A_8. The "delays" found in these experiments were considerably smaller than those expected theoretically. This may be due to the fact that the measurement of very small delays is necessarily somewhat inaccurate with the apparatus used, since the speed of the piston falls off as the compression approaches its maximum. In such cases the lowest ignition temperature is, of course, reached *before* the piston reaches the top of compression, so that the measured "delays" which are measured from the time of maximum compression tend to be too small. But there is also a curious effect, which is invariably observed in these experiments on the self-ignition of carbon compounds, when the initial temperature is high, and the time of explosion short. It is always found that the explosion, though apparently sharp, is not complete, but that a fluffy deposit of carbon is thrown down. This deposition of carbon in an explosion has often been noticed by other workers when ignition is effected by a spark, but it is usually thought to be a consequence of having too little oxygen for complete combustion; in our experiments, however, the oxygen was always in considerable excess. When the minimum ignition temperature is not greatly exceeded, and when therefore the explosion is comparatively slow, combustion is complete, and no carbon deposit is formed. At higher initial temperature, however, one cannot escape the conclusion that the hydrogen is burnt preferentially to the carbon, and that the rate of combination of carbon atoms can be greater than the rate of combination

of carbon with oxygen. The exact conditions when this occurs seems well worthy of further investigation.

It is always necessary to clean out the cylinder carefully after such a deposit has been formed, and before the next experiment is made; for, if not, abnormal results will be obtained, and the "delay" before ignition occurs will be very much shorter than is expected. To explain this it does not appear to be necessary to attribute any special "catalytic" activity to the carbon; a simple physical explanation seems to be sufficient. Such a deposit is known to be a very bad conductor of heat. If left on the walls of the piston and cylinder, we shall therefore have, on the next compression, large portions of gas from which the heat cannot get away quickly. Hence the ignition temperature is lowered, and the explosion takes place more rapidly.

XIII. Experiments on the self-ignition of mixtures of the vapour of ethyl ether $C_2H_5.O.C_2H_5$ gave very similar results. The results of the experiments on the compression of non-turbulent mixtures are shown in Table VIII.; while

TABLE VIII.

Ignition of mixtures of ether and air. Fan stationary.
Initial pressure (atmospheric) = 14·77 lb./sq. in.
Mean apparent value of "γ" = 1·309.
Strength of mixture = 1 part ether to 15 of air by weight.

No. of expt.	Compn. ratio.	Initial temp.	Max. Compn. pressure.	Max. temp. calc. from $\gamma=1{\cdot}309$.	Delay obs.	Cooling factor $=a$.
C_1	4·51	25° C.	105 lb./sq. in.	201° C.	No. ign.	0·47
C_2	4·83	24	116 ,,	211	,,	0·47
C_3	5·02	23	121·5 ,,	214	0·41	—
C_4	5·21	23	128 ,,	220	0·30	—
C_5	5·42	23	135·5 ,,	226	0·205	—
C_6	5·63	23	143 ,,	232	0·12	—
C_7	5·82	23	149 ,,	237	0·093	—
C_8	6·03	23	155 ,,	242·5	0·071	—
C_9	6·56	23	172 ,,	256	0·035	—

Table X. shows the results when the fan was rotating at full speed. No ignition was observed in experiments C_1 and C_2 in the first series, and in experiments D_1, D_2, and D_7 in the second series. The details of the calculation of the cooling factor corresponding to experiment C_2 are shown in Table IX. The cooling curve was not so regular in the case of C_1, but the mean value of the constant was the same,

TABLE IX.

Calculation of cooling factor corresponding to C_2.

Time from maximum pressure.	Observed pressure.	$P-P_f$.	$\log\frac{P_{max.}-P_f}{P-P_f}=\delta$.	$\alpha=2{\cdot}3\,\delta/t$.
0·0	116·2	44·9	0	—
0·232 sec.	111·7	40·6	·044	·44
·465 „	107·3	36·0	·096	·47
·70 „	103·3	32·3	·143	·47
·93 „	100·6	29·3	·185	·46
1·16 „	96·9	25·6	·244	·48
1·39 „	94·7	23·4	·283	·47
1·63 „	92·5	21·2	·326	·46

$P_f=4{\cdot}83\times14{\cdot}77=71{\cdot}3$. Mean = 0·47

TABLE X.

Ignition of mixtures of ether and air. Fan full speed.

Initial pressure (atmospheric) = 14·77 lb./sq. in.
Mean apparent value of $\gamma=1{\cdot}308$.
Strength of mixture as in Table VIII.

No. of expt.	Compn. ratio.	Initial temp.	Max. Comp. pressure.	Max. temp. calc. from $\gamma=1{\cdot}308$.	Delay obs.	Cooling factor $=\alpha$.
D_1	5·42	24° C.	135	226·5° C.	No. ign.	1·33
D_2	5·73	24	146·5	235	„	1·28
D_3	6·03	24	155·5	243 5	0·10	—
D_4	6·33	24	164	251	0·047	—
D_5	6·63	24	174	259	0·038	—
D_6	7·02	24	186	268·5	0·026	—
D_7	5·82	22	149	235	No. ign.	1·25
D_8	5·94	22	150·5	238	0·13	—
D_9	6·03	22	155	240	0·087	—
D_{10}	6·13	22	161	243	0·078	—

Mean = 1·29

namely $\alpha=0{\cdot}47$. Table XI. shows the details of the calculations for D_2; it will be noticed that the cooling factor appears to diminish fairly steadily as the time increases. D_1 and D_7 showed a similar effect; in each case the mean has been taken, and the mean value for the three experiments, namely $\alpha=1{\cdot}29$, is probably fairly accurate. The

values for the cooling factors so obtained are lower than those found in the heptane experiments; this would be expected, for, although the specific heats of the mixtures are about the same, the compression ratios used were lower in the case of ether mixture than in the case of heptane, since ether has a considerably lower ignition temperature.

TABLE XI.

Calculation of cooling factor corresponding to D_2.

Time from maximum pressure.	Observed pressure.	$P-P_f$.	$\log \frac{P_{max.}-P_f}{P-P_f}=\delta$.	$a=2{\cdot}3\,\delta/t$.
0	146·5 lb./sq. in.	61·8	—	—
·232 sec.	129·5 „	44·8	·140	1·39
·465 „	118·4 „	33·7	·263	1·30
·70 „	109·5 „	25·8	·379	1·25
·93 „	103·6 „	19·9	·492	1·22
1·16 „	97·7 „	13·0	·677	1·34
1·39 „	96·2 „	11·5	·730	1·21
$P_f=5{\cdot}73\times14{\cdot}77=84{\cdot}7$ lb./sq. in.				Mean=1·28

XIV. Taking the ignition temperature T_0 (without fan) as 212° C.=485 abs., the experimental observations of the delay before ignition at higher temperatures are reproduced closely if the value of $\frac{b}{T_0}=11$. In this case, we have

$$a\frac{T_0-\theta}{T_0}=0{\cdot}47\times\left(\frac{212-23}{485}\right)=0{\cdot}183.$$

The theoretical values for "t," calculated according to equation (11), are shown in the following table.

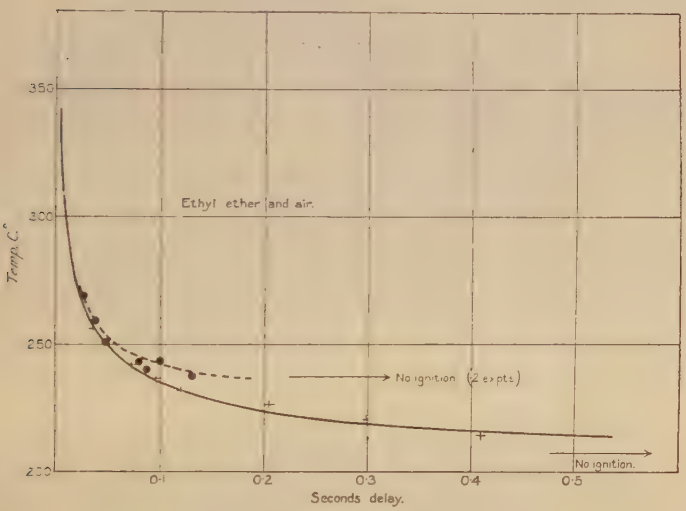
TABLE XII.

T/T_0.	T.	t° C.	$a\left(\frac{T_0-\theta}{T_0}\right)t$ (theor.).	t (calc.).
1·004	487	214	·098	0·535 sec.
1·01	490	217	·065	·355
1·03	499·5	226·5	·030	·164
1·07	519	246	·0105	·057
1·11	538·5	265·5	·0046	·025
1·27	616	343	·0004	·002

The lower curve in fig. 10 is the theoretical curve derived in this way; the experimental points are shown by crosses.

In the series of experiments with the fan running at full speed, "α" = 1·29, and the new ignition temperature T_0' is approximately 235° C. = 508 abs.

Fig. 10.



Hence $$\alpha\left(\frac{T_0'-\theta}{T_0'}\right)=1{\cdot}29\times\left(\frac{235-23}{508}\right)=0{\cdot}54,$$

and $$\frac{Q_0'}{Q_0}=\frac{1{\cdot}29\,(235-23)}{0{\cdot}47\,(212-23)}=3{\cdot}08.$$

Hence if $\frac{b}{T_0}=11$, T_0' should be given by

$$\log 3{\cdot}08=0{\cdot}489=11\left(1-\frac{T_0}{T_0'}\right),$$

or $$\frac{T_0}{T_0'}=0{\cdot}956\,;\qquad\therefore\quad T_0'=507{\cdot}5\ \text{(abs.)},$$

which is evidently very close to the observed value.

Taking this value for T_0', we have

$$\frac{b}{T_0'}=11\times\frac{485}{507{\cdot}5}=10{\cdot}5.$$

Table XIII. gives the theoretical values for the time of ignition at various temperatures, calculated as already described, taking $T_0' = 507{\cdot}5$; $\frac{b}{T_0'} = 10{\cdot}5$.

Table XIII.

T/T_0.	T.	t° C.	$a\left(\frac{T_0'-\theta}{T_0'}\right)t$ (theor.).	t (calc.).
1·004	509·5	236·5	0·104	0·19 sec.
1·01	512·5	239·5	0·069	0·13 „
1·03	522·5	249·5	0·0325	0 06 „
1·07	543	270	0·012	0·022 „
1·11	563·5	290·5	0·0055	0·010 „

The corresponding curve is the upper dotted curve in fig. 10, the experimental values being shown by circles. The general agreement is again all that could be desired. Although the ignition temperature of ether is very much lower than that of heptane, the temperature coefficient of the combustion reaction is the same within the experimental errors involved.

XV. The experiments on carbon bisulphide were expected to be of particular interest, on account of the anomalous behaviour of this substance if used as a fuel in internal combustion engines. It is known that for any given fuel the highest thermal efficiency obtainable is limited mainly by the tendency to "detonation" at high compression ratios. It is usually assumed that the tendency of any fuel to detonate depends upon its ignition temperature; the lower the ignition temperature, the greater will be the tendency to detonation. According to our views, the ignition temperature is not a safe criterion of the tendency to detonation; the temperature coefficient is also an important factor which must be taken into account. The use of carbon bisulphide as a fuel illustrates this point very well; although it has a lower ignition temperature than heptane, it detonates less easily in internal combustion engines, and not more easily, as might be supposed. We expected, therefore, to find, by the experiments described in this paper, that the temperature coefficient of its reaction with oxygen was very distinctly lower than that of heptane and similar substances. The experiments fully confirmed this, although the results do not appear to be so satisfactory in all respects as those carried out with heptane and ether.

Tables XIV. and XV. contain all the experimental results of the ignition at various temperatures of mixtures of CS_2 and air (*a*) when initially stagnant, and (*b*) when the fan was rotating at full speed.

TABLE XIV.

Ignition of mixtures of CS_2 and air. Fan stationary.
Initial pressure (atmospheric) = 14·60 lb./sq. in.
Mean apparent value of "γ" = 1·332.
Strength of mixture = 1 part CS_2 to 8 parts air by weight.

No. of exp.	Compn. ratio.	Initial temp.	Max. Compn. pressure.	Max. temp. calc. from $\gamma=1{\cdot}332$.	Delay obs.	Cooling factor.
E_1 ...	6·03	49·5° C.	159 lb./sq. in.	313° C.	0·086	
E_2 ...	5·63	47·5	143 „	296	0·115	
E_3 ...	5·63	42	143 „	286	0·18	
E_4 ...	5·63	39	147 „	281	0·26	
E_5 ...	5·63	36	147 „	275·5	0·35	
E_6 ...	5·63	32	145 „	268·5	0·42	
E_7 ...	5·02	32	126·5 „	248	No ign.	0·46
E_8 ...	5·21	31·5	132·5 „	254	0·71	
E_9 ...	5·42	31·5	140·5 „	261	0·59	
E_{10}...	5·82	31·5	154·5 „	273·5	0·34	
E_{11}...	7·02	31	193 „	308	0·087	

TABLE XV.

Ignition of mixtures of CS_2 and air. Fan full speed.
Initial pressure (atmospheric) = 14·6 lb./sq. in. expts. F_1–F_6.
= 14·5 „ „ F_7–F_{10}.
Mean apparent value of $\gamma = 1{\cdot}323$.
Strength of mixture as before.

No. of expt.	Compn. ratio.	Initial temp.	Max. Compn. pressure.	Max. temp. calc. from $\gamma=1{\cdot}323$.	Delay obs.	Cooling factor.
F_1 ...	6·33	43°C.	169 lb./sq. in.	300° C.	No ign.	1·30
F_2 ...	6·44	43	171 „	304	0·14	—
F_3 ...	6·56	43	177 „	307	No ign.(?)	1·38
F_4 ...	6·63	42·5	178 „	309	0·13	—
F_5 ...	6·84	43	185 „	315	0·12	—
F_6 ...	7·55	43	208 „	334	0·05	—
F_7 ...	6·56	43	175 „	307	0·11	—
F_8 ...	7·02	43	191 „	320	0·09	—
F_9 ...	8·02	43	224 „	346	0·03	—
F_{10}...	10·0	42	? „	391	Ignit'd before top of compn.	—

Mean = 1·34

The cooling curves were very uniform in the above cases.

XVI. It is clear, from the shape of the curve connecting time of ignition with temperature, from the results of experiments without the fan, and from the difference in ignition temperature observed with and without the fan, that the temperature coefficient of the reaction is low. If we take T_0, the ignition temperature without the fan, to be 253° C. = 526 abs., and $\frac{b}{T_0} = 7$, we obtain the theoretical values for the time of ignition from equation (11) which are given in Table XVI.

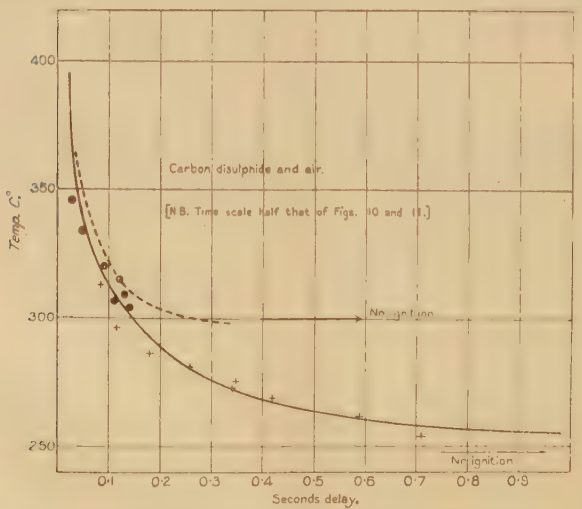
The value of $\alpha\frac{T_0-\theta}{T_0}$ is $0{\cdot}46 \times \frac{253-32}{526} = 0{\cdot}193$.

TABLE XVI.

$T_0 = 526$ abs. $b/T_0 = 7{\cdot}0$.

T/T_0.	T (abs.).	θ° C.	$\alpha\left(\frac{T_0-\theta}{T_0}\right)t$ (theor.).	t (calc.).
1·004	5228	255	0·187	·98 sec.
1·01	531	258	0·133	·70 „
1·03	542	269	0·072	·38 „
1·07	563	290	0·036	·19 „
1·11	584	311	0·020	·105 „
1·27	668	395	0·004	·021 „

Fig. 11.



The theoretical values are shown by the lower curve in fig. 11.

The cooling factor when the fan is used is 1·34. The ratio of the loss of heat (or reaction velocity) at the ignition temperatures with and without fan is consequently

$$\frac{Q_0'}{Q_0} = \frac{1{\cdot}34\,(296-42)}{0{\cdot}46\,(253-32)}$$
$$= 3{\cdot}35.$$

Hence the new ignition temperature T_0' should be given by the expression

$$\log 3{\cdot}35 = 0{\cdot}525 = b/T_0\left(1 - \frac{T_0}{T_0'}\right)$$
$$= 7\left(1-\frac{T_0}{T_0'}\right);$$

or
$$\frac{T_0}{T_0'} = 1 - 0{\cdot}075$$
$$= 0{\cdot}925,$$
$$T_0' = 568{\cdot}5 = 295°{\cdot}5\ \text{C.}$$

The new value of $\frac{b}{T_0'}$ corresponding to this is therefore

$$\frac{b}{T_0'} = 7 \times \left(\frac{526}{568{\cdot}5}\right) = 6{\cdot}5.$$

The calculation of the corresponding delay curve is shown in Table XVII., the value of $\alpha\left(\frac{T_0'-\theta}{T_0'}\right)$ being

$$1{\cdot}34 \times \frac{254}{569} = 0{\cdot}6.$$

Table XVII.

T/T_0.	T.	t° C.	$\alpha\left(\frac{T_0'-\theta}{T_0'}\right)t$ (theor.).	t (calc.).
1·004	571	298	0·205	·34 sec.
1·01	574	301	0·15	·25 ,,
1·03	585·5	312·5	0·08	·13 ,,
1·07	608·5	335·5	0·042	·07 ,,
1·11	631	358	0·025	·04 ,,

This is the dotted curve shown in fig. 11, the experimental points being marked as before. The agreement between

theory and experiment is in this case only moderate. In particular, it will be noticed that although the ignition temperature with the fan running is calculated, from the results without the fan, to be 295°·5 C., actually no ignition was observed at 300°C., and even in one experiment at 307°C., although this is extremely doubtful, since in two other experiments ignition was observed to take place at 307° C. and 304° C., with the comparatively short delays of 0·11 and 0·14 second respectively. It is possible that when the loss of heat is considerable, *and* the temperature coefficient small, there is an appreciable error introduced in neglecting changes of concentration when calculating the time of ignition at temperatures near the ignition temperature. This would account for no ignition being observed when a long delay was expected. The value of b/T_0 given above cannot, however, be very far wrong. For suppose we take the ignition temperature T_0' of the mixture when the fan is running as 300° C. instead of 295°·5, and calculate the temperature coefficient solely from the difference in T_0 and T_0' without regard to the "delay" curve when the fan is stationary; then we shall have

$$\frac{T_0}{T_0'} = \frac{526}{573} = 0{\cdot}918$$

and
$$\frac{Q_0'}{Q_0} = \frac{1{\cdot}34\,(300-42)}{0{\cdot}46\,(253-32)} = 3{\cdot}40\,;$$

$$\therefore \quad \log 3{\cdot}40 = {\cdot}531 = \frac{b}{T_0}\left(1-\frac{T_0}{T_0'}\right)$$

$$= \frac{b}{T_0} \times {\cdot}082\,;$$

$$\therefore \quad \frac{b}{T_0} = 6{\cdot}5 \text{ instead of } 7{\cdot}0.$$

This value for $\frac{b}{T_0}$ would, in fact, fit the lower part of the delay curve without the fan rather better than the value 7·0, but the calculated delay curve when the fan is running would then be some way from the experimental points. If we take a mean value

$$b/T_0 = 6{\cdot}7,$$

we shall be very unlikely to be as much as 10 per cent. from

the true value, even when all possible sources of error are taken fully into account.

XVII. It has already been mentioned that when the time of ignition by compression of hydrocarbons (and of ether) is small, *i. e.* when the gases are suddenly compressed to a temperature well above the ignition temperature, carbon is thrown down, even though excess of oxygen is present. In the experiments with CS_2 an even more curious phenomenon was noticed. In this case, whereas the sulphur burns to SO_2 when the initial temperature of compression does not exceed very greatly the lowest ignition temperature, when the initial temperature is high the products of combustion smell strongly of H_2S. For instance, the products of combustion in experiments E_1 and E_{11} above both smelt strongly of H_2S, although in E_2–E_{10} inclusive only SO_2 could be detected by smell. It was also possible to detect H_2S after experiments F_6 and F_{10} (with the fan), the smell being particularly strong in the case of F_{10}. The smell of H_2S could also just be detected along with SO_2 in experiment F_9, whereas in the remainder only SO_2 could be detected.

The H_2S could only have come by combination with water-vapour present in the air, which was not dried. This occurrence of H_2S is all the more interesting since it is known that a perfectly dry mixture of CS_2 and oxygen can be exploded by a spark, whereas perfectly dry mixtures of other gases, *e. g.* carbon monoxide with oxygen, cannot. It is possible that some such reaction as

$$CS_2 + 2H_2O = CO_2 + 2H_2S$$

takes place, followed by the combustion of H_2S ; but even if this is the case, it would be expected that the H_2S would be quickly burnt in presence of excess of oxygen under the conditions of these experiments. Further experiments on the ignition of H_2S itself will probably throw some light on these observations.

XVIII. In Table XVIII. are summarized the chief results of the experiments described above. The ignition temperatures represent the lowest average temperatures at which the non-turbulent mixture could be caused to ignite. The rates of evolution of heat at these temperatures for the three cases are calculated from the cooling factors and the specific heats of the mixtures.

TABLE XVIII.

	Heptane C_7H_{16}.	Ether $C_2H_5.O.C_2H_5$.	Carbon bisulphide CS_2.
Composition of gas by weight...	1 : 20 of air.	1 : 15	1 : 8
T_0=ignition temperature	280° C.	212° C.	253° C.
c_v at ignition temperature	0·20	0·20	0·18
Rate of evolution of heat due to reaction per gram of mixture	25 calories per second.	18	18·5
Total heat of combustion per gram of mixture	510 calories.	510 calories.	386 calories.
Value of b/T_0	10·0±5 %	11·0±5 %	6·7±10 %

XIX. In order to calculate the true temperature coefficient B (see equation 5) from the values of b/T_0, it is necessary to examine the significance of T_0 a little more closely. As already stated, T_0 is a measure of the lowest average temperature of the gas at which ignition takes place. Now the actual temperature of the gas after sudden compression can hardly be uniform throughout : in fact, when the gas ignites after a considerable delay, it is always found that the pressure, and therefore the average temperature, falls, in some cases quite considerably, before ignition takes place throughout the mass. This shows clearly that that portion of the gas which ignites at first has initially a higher temperature than the average, thus confirming Dixon's experiments. Absence of information as to the temperature gradients which may exist under these conditions has no doubt led Nernst and Dixon in their experiments to calculate the ignition temperature as if the compression were adiabatic, and to ignore the influence of loss of heat during compression and before ignition. They assume, in fact, that that portion of the gas which does ignite is at the adiabatic temperature.

It is hardly likely, however, that big differences in temperature exist after compression when the gases are in a turbulent state ; and the fact that the temperature coefficients, calculated from the differences in "average" ignition temperature between turbulent and non-turbulent mixture,

agree so well with those calculated from measurements of time of ignition at various temperatures with non-turbulent mixture, confirms the views taken in the previous paper (Tizard, *loc. cit.*) that it is unlikely also that any big differences of temperature exist in the non-turbulent mixture after compression. In the absence of direct evidence on this point, however, it is important to calculate the "adiabatic" temperatures also in the above cases.

The mean specific heats C_v per gram molecule are :

Heptane (room temperature–300° C.) = 50 calories approx.

Ether (65°–230° C.) = 33·6 calories (Regnault).

CS_2 (70°–194° C.) = 10·0 (Regnault).

Taking these values, and $C_v = 5·0$ for air, we obtain the figures for the mean true value of "γ," and the corresponding adiabatic temperatures, given in the following table.

TABLE XIX.

	Heptane.	Ether.	CS_2.
Mean apparent value of "γ" (observed)	1·313	1·309	1·332
Mean true value of γ	1·353	1·347	1·384
"Average" ignition temperature	553 abs.	485 abs.	526 abs.
"Adiabatic" ignition temperature	594	516	572
Difference	41° C. = 7·4 %	31° C. = 6·4 %	46° C. = 8·7 %

The average specific heat for CS_2 taken in the above calculations is probably too low, since it refers only to a range of temperature up to 194° C., whereas the ignition temperature was 250° C.

It will be observed that the difference between the average observed and the theoretical adiabatic temperatures is not very great; we consider that the "average" temperature is probably closer to the true ignition temperature than is the "adiabatic" temperature, but for the purpose of estimating every possible source of error in the temperature coefficients, it is better at this stage to recognize the uncertainty, and take for the true values of the ignition temperatures the values

Heptane 573° absolute
Ether 500° ,,
CS_2 549° ,,
} ± 4 per cent.

Hence, from the values of $\frac{b}{T_0}$ we get finally for the true temperature coefficient B :

	Temperature coefficient B (equation 5).
Heptane-air...	13,200 ± 9 per cent.
Ether- „ ...	12,600 ± 9 „
CS_2- „ ...	8,500 ± 14 „

The significance of these figures will perhaps be better appreciated by the statement, that the percentage rise in absolute temperature necessary to treble the reaction velocity is 4 per cent. in the case of heptane and ether and 7 per cent. in the case of CS_2.

XX. Of recent years, considerable attention has been directed to the "radiation" theory of chemical reactions. According to this theory, the ultimate cause of any chemical reaction is to be found in the absorption of radiation of a frequency which depends upon the nature of the reactants. In the case of the majority of chemical reactions, namely those which are not "photochemical" in nature, this radiation will belong either to the visible, or more usually in the short infra-red part of the spectrum. The supporters of the theory hold the view that it is only through the absorption of such radiation that a molecule is able to acquire that excess of energy, over the average at any temperature, which enables it to decompose or to react with another molecule. The frequency of the radiation is therefore known as the "activating" frequency.

This reasoning leads to the conclusion that the temperature coefficient B of a mono-molecular reaction is determined by the relation

$$B = \frac{h\nu}{k}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where "ν" is the activating frequency, which should correspond to an absorption band in the reacting species. No reliable experimental evidence has yet been brought forward in support of this theory, but in view of the scantiness of the data existing on homogeneous gas reactions, it is of particular interest to apply it to the results of the experiments described above.

In attempting to apply the theory, a difficulty at once arises. Evidence has been brought forward to show that the ignition temperature of substances with oxygen is

practically independent of the concentration of the combustible substance. If the rate of the reaction were determined solely by the amount of oxygen present, we might expect the temperature coefficient also to depend solely on the oxygen, and therefore to be the same in all cases. This is clearly not true. Nor does the temperature coefficient, in the case of the heptane explosion which has been most closely investigated, correspond at all closely to that calculated by means of equation (12) from the infra-red absorption of oxygen. Oxygen has an absorption band corresponding to $\lambda = 3{\cdot}2\ \mu$, or $\nu = {\cdot}94 \times 10^{14}$; hence, since $\frac{h}{k} = 4{\cdot}86 \times 10^{-11}$, we should have

$B = 4550$ (calc.) instead of $13{,}200 \pm 9$ per cent. (obs.).

It is clear that equation (12) cannot be applied. On the other hand, if the rate of reaction depended on the product of the number of active molecules, both of oxygen and the other reactant, we should expect, on the same theory, to find the temperature coefficient given by

$$B = \frac{h}{k}(\nu_1 + \nu_2),$$

where ν_1, ν_2 correspond to absorption bands in the reacting substances.

Now, all hydrocarbons have a weak absorption band at $\lambda = 2{\cdot}4\ \mu$, and a fairly strong one at $3{\cdot}43\ \mu$. Taking $\lambda = 2{\cdot}4\ \mu$, which is most favourable to the theory, we have $\nu_2 = 1{\cdot}25 \times 10^{14}$ and ν_1 (oxygen) $= {\cdot}94 \times 10^{14}$.

Hence $B = 4{\cdot}86 \times 10^{-11} \times 2{\cdot}2 \times 10^{14} = 10{,}700$.

This approaches more closely the experimental value

$$B = 13{,}200 \pm 9 \text{ per cent.}$$

It must be pointed out, however, that this approximate agreement is only obtained by an assumption as to the actual mechanism of the reaction which does not agree with the existing experimental results.

The failure of the "radiation theory" to account for the results obtained in these experiments is more significant when we regard it in a different way. The theory requires that the rate of a chemical reaction should be proportional to the density in the reacting system of the radiation which is absorbed by the reacting substances. Now, in the case of gases which are caused to react by a rise in temperature due to sudden compression, the radiation *density* must remain practically unchanged, for the temperature of the walls remains constant. It may be momentarily increased

owing to the sudden compression, but such an increase cannot persist during the period of delay, and in any case is negligible compared with the increase in radiation density which would occur if the temperature of the *walls* of the vessel were raised to the compression temperature of the gas. Again, the emission and absorption of radiation by the gas itself at the compression temperature of 500–600 absolute is negligible compared with that of the solid walls. Hence we arrive at the conclusion that, although the density of radiation in the system is not appreciably changed, the gases react at a high rate. This fact appears to us to prove conclusively that the radiation theory cannot be accepted either in its original form, or as modified to meet its failure to account quantitatively for the temperature coefficients of chemical reactions in liquids under steady conditions of temperature.

It must be pointed out, however, that in spite of this conclusion, there does seem to be an indirect connexion between the effect of temperature on the rate of combustion of many substances and their absorption of infra-red radiation. For example, Coblentz has shown that all paraffin hydrocarbons have very similar absorption spectra, with a weak band at $\lambda =$ about 2·4 μ, and strong bands at $\lambda = 3{\cdot}43$, 6·86 μ, etc. Now, Ricardo has shown that the tendency of hydrocarbon fuels to detonate in an internal combustion engine depends consistently on their ignition temperatures as determined in the manner described above. According to our views this is strong evidence that the temperature coefficients are practically the same throughout. Again, it has been shown that ethyl ether has approximately the same temperature coefficient of combustion as that of heptane; while Coblentz has found that its absorption spectrum is also nearly identical, with bands at 2·4 μ and 3·45 μ. Carbon bisulphide, on the other hand, has a much smaller tendency to detonate in an internal combustion engine than heptane, although it has a lower ignition temperature; corresponding to this we find that the temperature coefficient is low, and that the first strong absorption band in the infra-red occurs at $\lambda = 4{\cdot}6$ μ. Finally, hydrogen "detonates" easily in spite of its high ignition temperature; its temperature coefficient must therefore also be high, a deduction which is confirmed by some preliminary experiments we have made on the delay before the ignition of a non-turbulent mixture of hydrogen and air. We should expect from this point of view to find an absorption band in the short infra-red region (say about 1·0 μ); actually no absorption is observed, but that the

frequency of atomic oscillation is high is in agreement with our general knowledge of the hydrogen molecule. In spite, therefore, of the strong arguments that have been brought forward against the radiation theory of chemical reactions, these results support the view that there is a connexion, even though an indirect one, between the temperature coefficients of gaseous reactions and the infra-red spectra of the reacting substances.

XXI. The results of this investigation may be summarized as follows:—

(*a*) Quantitative experiments confirm the view that at the lowest ignition temperature the heat evolved by the combustion of a gas just exceeds that lost to the surroundings.

(*b*) From measurements of the rate of loss of heat just below the ignition temperature, and of the intervals between the end of compression and the occurrence of ignition at different temperatures, it is possible to deduce the temperature coefficient of the gaseous reaction.

(*c*) The temperature coefficients so obtained are confirmed by the increase in the minimum ignition temperature which is observed when the gas is in a turbulent state.

(*d*) The results show that the temperature coefficient of the combustion of carbon bisulphide is much lower than that of heptane or ether. This is in agreement with the relative tendencies of these fuels to detonate in an internal combustion engine.

(*e*) The results do not agree with the radiation theory of chemical reaction.

(*f*) Some evidence is put forward to show that the rate of reaction on sudden compression is independent within wide limits of the concentration of the combustible gas, but only depends on the amount of oxygen present. This evidence is, however, incomplete.

We are greatly indebted to Messrs. Ricardo & Co. for the loan of their apparatus and for much additional assistance; also to the Department of Scientific and Industrial Research for a grant towards the expenses of the investigation. We also take this opportunity of thanking Mr. C. T. Travers for his help in carrying out some of the experiments.

IX. *On the Vibration and Critical Speeds of Rotors.* *By* C. Rodgers, *O.B.E., B.Sc., B.Eng., M.I.E.E.**

NUMEROUS papers have been written on the question of the whirling and vibration of loaded shafts and kindred subjects, and the calculation of the first critical speed—the lowest speed at which the vibration shows a maximum value, is now a matter of daily routine in designing offices.

This critical speed can be calculated with sufficient accuracy for practical purposes and as a rule the running at speeds not in the neighbourhood of that indicated by the calculation is free from vibration. But cases occasionally arise where troubles from vibration occur at speeds above or below the calculated critical speed, the reason for which is obscure and the remedy correspondingly difficult to find.

It is the object of this paper to discuss various subsidiary causes which might conceivably lead to unsatisfactory running at other than the usual calculated critical speed, but while these are indicated as possible causes of disturbance, it is not to be assumed that these causes always exist or that they will always induce disturbed running. The object is rather to indicate reasons why vibration might possibly arise and thus if an actual case occurs, to suggest a clue to the cause.

Although the basis of the paper is a physical or mechanical one, the treatment is largely mathematical, as it is only by this means that formulæ can be obtained from which numerical results can be worked out.

The phenomena when a rotor vibrates are complicated, as the shaft is supported in the bearings on a film of oil, the thickness of which is continually changing, the bearings and foundations are not themselves perfectly rigid, and there is a certain amount of initial bending of the shaft (and to a much smaller extent of the rotor body) due to gravity. If the rotor consists of a number of disks as in a steam turbine, there is also the inter-action of the forces of each disk on the others.

For the sake of simplicity, we shall confine our attention to the case of a single part rotor, either a disk or a cylinder, rigid as regards bending and mounted on an elastic shaft running in rigid bearings. Some effects of non-rigidity of the rotor and bearings and of alterations in the thickness of the oil film in the bearings will be indicated.

* Communicated by the Author.

A single part rotor can vibrate in either of two ways, as shown in figures 1 and 2, or in a manner which is a combination of the two motions :—

Fig. 1.

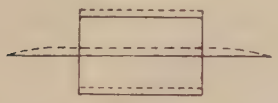


Fig. 2.

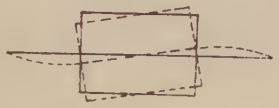


Fig. 1 shows a purely transverse vibration, while in fig. 2 the motion is solely one of oscillation about the centre of gravity. In the transverse vibration the conditions are clearly the same whether the rotor body is a disk or is cylindrical ; but in the case of the oscillation, the motion, owing to the gyrostatic effects called into play, depends both on the proportions of the rotor and the speed at which it is running.

The speed at which transverse vibration becomes a maximum we shall call the "first critical speed," and that at which the oscillation becomes a maximum, the "second critical speed," as the latter is in all practical cases considerably higher than the former.

The following is a general outline of the treatment adopted and the conclusions reached :—

Section 1 deals with the vibration of a rotor when not running, and a relationship is deduced between the frequencies for the transverse motion and for the oscillation which we shall call respectively the "stationary first critical speed," and the "stationary second critical speeds."

The second section deals with the transverse vibration, frictional resistance being ignored. It is first shown that there appears to be no foundation for the frequently expressed view that there is a possible region of marked vibration at $\frac{1}{\sqrt{2}}$ times the first critical speed, as such a conclusion can only be reached through an incorrect assumption with regard to the conditions. It is then shown that the motion or vibration is a circular whirl about the statically deflected position of the shaft, and that this motion reaches a maximum at a speed equal to the stationary first critical speed. The magnitude of the whirl is proportional to the amount by which the machine is out of balance, so that the main vibration here dealt with should disappear with good balancing.

The action of gravity is then gone into more fully, and it

is shown that in addition to producing the ordinary static deflexion, the action of gravity is such as to cause a double frequency ripple in the whirl which would tend to reach a maximum at half the first critical speed. The magnitude of this ripple is, however, proportional to the square of the amount by which the rotor is out of balance, and would therefore fail to appear in a well-balanced machine. In any case the effect is very small.

It is then shown that a rotor with bi-polar asymmetry, such as exists in a rotor slotted for a two-pole winding, may show a double frequency vibration at half the critical speed even when the rotor is perfectly balanced, so that such a machine might vibrate at half the critical speed even when it would run perfectly at the full critical speed. Vibration arising from this cause could not, therefore, be rectified by balancing, and this is the only case met with where improved balancing would not effect an improvement in the running. This case is gone into in some detail, and it is shown that the motion here also is a circular whirl of double frequency, that is, of twice the speed of rotation of the machine. If, in addition, the machine is out of balance, a triple frequency effect might appear, but is not likely to do so.

The effect is then discussed of lack of proportionality in the deflexion of the shaft and again the possibility appears of vibration appearing at half the critical speed, but only if the machine is not properly balanced. The effect is then gone into of fluctuations in the angular velocity through variations in the driving torque, and of resonance between the rotor and the foundations or other masses outside the machine, from which it appears that marked vibration might appear at almost any speed through either of these causes.

The effect of friction on the transverse vibration is then discussed, and the results are given for the case where the frictional resistance varies as the first power of the speed, and also where it varies as the second power of the speed, the latter being more probably in accordance with the facts than the former. It is shown that the maximum vibration appears in both cases at a speed equal to the stationary critical speed, also that the phase difference between the force due to the out-of-balance and the displacement depends on the amount of friction, and also on the speed. If the frictional forces vary as the square of the speed, as is probably the case, the angle varies also with the amount by which the rotor is out-of-balance.

Some effects of viscosity of oil in the bearings, and of different bearing clearances are then gone into.

In Section 3, the oscillatory vibration is dealt with, taking into account the gyrostatic effects when the machine is rotating, but ignoring the friction in order to keep the expressions as simple as possible. It is there shown that the gyrostatic effect causes the point of marked vibration to occur at a higher speed than would be the case if the machine were not rotating, and simple rules are given for calculating this vibrating speed. An example is added to illustrate the method of working the rules given.

Much of the work on the main transverse vibration and the main oscillatory vibration has been dealt with in various forms by Chree, Stodola, Morley and others, and the solution for the transverse vibration with friction depending on the first power of the speed has been given by H. H. Jeffcott (Phil. Mag. March 1919), but the ground covered by the remainder of the paper, particularly the question of subsidiary critical speeds, does not appear to have received much attention; there is, however, in 'Engineering' a discussion where subsidiary critical speeds are touched on, arising out of a paper by W. Kerr in that journal (Feb. 18th, 1916).

Section I.—Stationary Vibrations.

A. *Transverse Vibrations.*

1. If M is the mass of the rotor body (the mass of the shaft being being neglected), and we assume the rotor to be perfectly balanced, the shaft will, when not rotating, show a deflexion measured at the centre of gravity of the rotor of

$$y_0 = -\frac{Mg}{\sigma}, \quad \ldots \ldots \quad (1)$$

where σ is the force required to produce unit deflexion. The method of working out the static deflexion of the rotor for actual cases is well understood and the value of σ can be got from the deflexion diagram.

2. If now the rotor is set in vibration in a vertical plane, the motion is represented by the following equation (using fluxional notation, where $\ddot{y}$ is written for $\frac{d^2y}{dt^2}$ and $\dot{y}$ for $\frac{dy}{dt}$, etc.)

$$M\ddot{y} + \sigma y + Mg = 0. \quad \ldots \ldots \quad (2)$$

The solution is

$$y = N_1 \sin\left(\sqrt{\frac{\sigma}{M}}t - \gamma_1\right) - \frac{Mg}{\sigma}, \quad . \quad . \quad . \quad (3)$$

where N_1 and γ_1 are constants the values of which depend on the initial conditions. The vibration therefore takes place about the statically deflected position as a centre, and with a frequency of vibration of $\frac{1}{2\pi}c_1$, where

$$c_1 = \sqrt{\frac{\sigma}{M}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

This vibration takes place in a vertical plane and may be considered as the resultant of two vectors rotating in opposite directions, each with an angular velocity of $\sqrt{\frac{\sigma}{M}}$. If σ and M are expressed in c.g s. or f.p.s. units, this angular velocity will be in radians per second and since from (1), $\frac{\sigma}{M}$ is numerically equal to $\frac{g}{y_0}$, the speed of either of these vectors in revs. per minute will be $\frac{60}{2\pi}\sqrt{\frac{g}{y_0}}$. If, further, g and y_0 are in c.g.s. units we have the formula

$$\text{R.P.M.} = \frac{60}{2\pi}\sqrt{\frac{981}{y_0}} = \frac{300}{\sqrt{y_0}} \quad \text{approximately,} \quad . \quad (5)$$

where y_0 is the static deflexion in cm.

3. It will be seen afterwards that, as is well known, this formula gives the first critical speed in R.P.M. ; this is to be expected, as the out-of-balance forces will then resonate with the natural free vibrations, with the result that the latter will become of considerable magnitude.

B. *Oscillatory Vibrations.*

1. If the rotor is twisted about its centre of gravity so that the deflexion is in a vertical plane, and is allowed to oscillate freely, the motion is represented by

$$B\ddot{\psi} + \kappa\psi = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where B is the cross moment of inertia, that is, the moment of inertia of the rotor about a line through the centre of gravity at right angles to the shaft, ψ is the angle through

which the axis of the rotor at its centre of gravity is deflected from the stationary position, and κ is the torque required to produce unit angular deflexion.

The solution is

$$\psi = N_2 \sin\left(\sqrt{\frac{\kappa}{B}}t - \gamma_2\right), \quad . \quad . \quad . \quad . \quad (7)$$

where N_2 and γ_2 are constants the values of which depend on the initial conditions. The frequency of the oscillation is therefore

$$\frac{1}{2\pi}\sqrt{\frac{\kappa}{B}} \text{ or } \frac{1}{2\pi}c_2, \text{ where}$$

$$c_2 = \sqrt{\frac{\kappa}{B}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

As already indicated, we cannot at once deduce from this what will be the actual second critical speed, owing to the gyrostatic effects, but the result is of importance, as it simplifies the calculation of the actual second critical speed, as will be shown later. We shall in what follows call c_2 the stationary second critical speed.

2. It should be pointed out that there is a simple relation between c_1 and c_2 which greatly facilitates the calculation of the stationary second critical speed in those cases where the centre of gravity of the rotor is midway between the bearings. If $2l$ is the distance between the bearing centres and P_2 the force exerted by the deflected shaft on either bearing,

$$\kappa\psi = 2P_2 l.$$

The angle ψ is very small so that the force P_2 is the same as would be required to depress the shaft through a distance ψl if the rotor were held rigidly. Now we have seen that the force Mg at the centre of gravity causes a transverse deflexion of $y_0 = -\frac{Mg}{\sigma}$, and as ψ is small,

$$y_0 = \psi l,$$

so that $Mg = \sigma\psi l,$

also $\kappa\psi = 2P_2 l$ and $P_2 = \frac{1}{2}Mg;$

therefore $\kappa\psi = Mgl,$

so that $\kappa = \sigma l^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$

and (8) becomes $c_2 = l\sqrt{\frac{\sigma}{B}}, \quad . \quad . \quad . \quad . \quad . \quad (10)$

and if $B = Mk_2^2$,

$$c_2 = \frac{l}{k_2}\sqrt{\frac{\sigma}{M}}, \quad \ldots \ldots \quad (11)$$

and comparing with (4) we thus get:

$$\frac{c_1}{c_2} = \frac{k_2}{l}. \quad \ldots \ldots \quad (12)$$

We thus find that

$$\frac{\text{First critical speed (transverse vibration)}}{\text{Stationary Second critical speed (oscillation)}} = \frac{\text{Radius of Gyration for the cross moment of inertia}}{\text{Half the distance between the bearing centres}}.$$

This is a useful formula for calculating the stationary second critical speed when the first is known, for cases where the centre of gravity is midway between the bearings.

It shows that with cylindrical rotors of this type the second critical speed must always be considerably above the first, and the only instance in normal designs in which the second critical speed could be lower than the first would be that of a flywheel mounted on a short shaft.

Section II.—Transverse Vibrations—First Critical Speed.

A. *Neglecting Frictional Resistance.*

1. It will simplify the treatment of this question if we first consider the case of a rotor unimpeded by frictional resistances set up by the air and then treat separately the effects produced by friction.

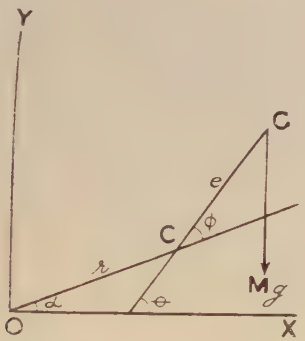
The conditions obtaining when a rotor is not perfectly balanced and is rotating are illustrated in fig. 3, where O represents the position of the centre line of the bearings, and C the deflected position of the centre line of the shaft, while G shows the position of the centre of gravity of the rotor. O thus gives the undeflected position of the shaft centre line and $OC = r$ the shaft deflexion at any instant, while $CG = e$ is the error in the centering of the rotor; Mg is the weight of the rotor acting vertically downwards.

The rotation of the rotor about its centre line, *i. e.* the rotation imparted by the prime mover, is represented by the motion of G around C, *i. e.* by the rate of change of θ.

The whirling of the rotor is represented by the motion of C about the undisturbed position of the shaft centre line, *i. e.* by the rate of change of a.

The "vibration" of the rotor is judged in a general way by the vibration of the bearings as felt when the hand is applied to them. The force on the bearings is that applied along OC by the deflexion r of the shaft, and vibration of

Fig. 3.



the bearings arises through the varying position and magnitude of OC; these in turn are due to the motion of the centre of gravity G.

2. If the machine is steadily rotating it might at first sight be thought that OC and CG would be in the same straight line, so that G would be steadily revolving together with C about the undisturbed position O of the shaft centre line with an angular velocity n say. At the same time the deflexion $OC=r$ might be changing its value and (neglecting the weight of the rotor) the motion would thus be given by

$$M\ddot{r}-Mn^2(r+e)+\sigma r=0,$$

or putting $\sigma/M=c_1^2$,

$$\ddot{r}+(c_1^2-n^2)r=n^2e,$$

the solution of which is

$$r=N_1\sin\left(\sqrt{\frac{1}{c_1^2-n^2}}t-\gamma_1\right)+\frac{n^2e}{c_1^2-n^2},$$

where N_1 and γ_1 are constants.

We should thus conclude that r would become unlimited

in magnitude when the angular velocity n is equal to c_1 or $\sqrt{\frac{\sigma}{M}}$, and, further, that the variation in the magnitude of r consists of a free vibration having a periodicity of $\frac{1}{2\pi}\sqrt{\frac{1}{c_1^2-n^2}}$.

This result for the periodicity of the free vibration would lead to the conclusions that when the machine is not rotating ($n=0$) the periodicity is c_1, the same as for the stationary transverse vibration, and that when running at the critical speed ($n=c_1$) the periodicity is zero. There would thus be some intermediate speed where the periodicity of the free vibration corresponds with the running speed, and resonance might take place. This would occur when

$$c_1^2-n^2=n^2$$

or

$$n=\frac{c_1}{\sqrt{2}}.$$

We should thus be led to expect marked vibration when the running speed is $\frac{1}{\sqrt{2}}\times$ the critical speed.

This conclusion and the argument on which it is based are, however, erroneous. In the first place, the assumption is made that ω or $\dot{\theta}$ is constant and further the condition is omitted that, as all the forces pass through O, the angular velocity about O must be constant, or $r^2\dot{\theta}=h$, say. The correct equations for the free vibrations are thus:

$$\left.\begin{aligned}\ddot{r}+(c_1^2-\dot{\theta}^2)r=0,\\ r^2\dot{\theta}=h.\end{aligned}\right\}$$

This does not admit of direct solution *, and it is simpler to use rectangular co-ordinates, as we shall now proceed to do.

* The solution is, however, well known and is given in books on Dynamics dealing with Central Forces:—

If p is the length of the perpendicular from the centre of force on the tangent to the path, it is known that

$$\frac{h^2}{p^3}\frac{dp}{dr}=\sigma r,$$

giving

$$\frac{h^2}{p^2}=\text{a constant}-\sigma r^2,$$

which is the pedal equation of a central ellipse.

3. With the notation given in fig. 3, it will be seen that the position of the rotor is completely defined by the co-ordinates x, y, and θ (x and y being the co-ordinates of G) and only three equations are required to express the motion fully: the value of the whirling angle α follows from the magnitude of the other co-ordinates.

The force exerted by the deflected shaft is σr, the components of which are

$$-\sigma(x-e\cos\theta) \text{ along OX and}$$

$$-\sigma(y-e\sin\theta) \text{ along OY.}$$

Resolving along OX and OY and taking moments about G, we thus have:

$$M\ddot{x}+\sigma(x-e\cos\theta)=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

$$M\ddot{y}+\sigma(y-e\sin\theta)+Mg=0, \quad . \quad . \quad . \quad . \quad (14)$$

$$Mk_1^2\ddot{\theta}+\sigma e(x\sin\theta-y\cos\theta)=0, \quad . \quad . \quad . \quad (15)$$

where k_1 is the radius of gyration of the rotor about the longitudinal axis through the centre of gravity.

In practice the rotor is driven at an average angular velocity ω say, which will vary from constancy only by small amounts which we shall find later are negligible.

Assuming as a first approximation that the angular velocity is constant $=\omega$, so that $\theta=\omega t$, (13) and (14) then become, writing $\sigma/M=c_1^2$ as before,

$$\ddot{x}+c_1^2x=c_1^2e\cos\omega t, \quad . \quad . \quad . \quad . \quad . \quad (16)$$

$$\ddot{y}+c_1^2y=c_1^2e\sin\omega t-g, \quad . \quad . \quad . \quad . \quad (17)$$

while (15) becomes an identity.

The solutions are

$$x=N_1\sin(c_1t-\gamma_1)+\frac{c_1^2e}{c_1^2-\omega^2}\cos\omega t, \quad . \quad . \quad . \quad (18)$$

$$y=N_2\cos(c_1t-\gamma_1)+\frac{c_1^2e}{c_1^2-\omega^2}\sin\omega t-g/c_1^2, \quad . \quad (19)$$

where N_1, N_2, and γ_1 are constants.

4. The above results give the motion of the centre of gravity G. The motion of the centre of the shaft C is to be found in a similar way, still on the assumption that the

angular velocity of rotation is constant. If x' y' are the co-ordinates of C,

$$x = x' + e\cos\omega t,$$

$$y = y' + e\sin\omega t,$$

giving the equations :

$$\ddot{x}' + c_1^2 x' = \omega^2 e\cos\omega t, \quad . \quad . \quad . \quad . \quad . \quad (20)$$

$$\ddot{y}' + c_1^2 y' = \omega^2 e\sin\omega t - g, \quad . \quad . \quad . \quad . \quad (21)$$

the solution of which is :

$$x' = N_1\sin(C_1 t - \gamma_1) + \frac{\omega^2 e}{c_1^2 - \omega^2}\cos\omega t, \quad . \quad . \quad . \quad (22)$$

$$y' = N_2\cos(C_1 t - \gamma_1) + \frac{\omega^2 t}{c_1^2 - \omega^2}\sin\omega t - g/c_1^2. \quad . \quad (23)$$

5. These equations are the same as for a perfectly balanced rotor with a weight attached to it of such small value as not to affect the position of the centre of gravity, or the downward pull due to gravity, but producing a force of $M\omega^2 e$. In other words, we can treat the unbalanced rotor as if it were a perfectly balanced rotor with a force $M\omega^2 e$ attached to it, and as this mode of presentation is easier to follow than the former, we shall employ it in the remainder of the paper.

6. It will be seen that the solutions for the motions of the centre of gravity and the centre line of the shaft are the same, except that the former has an amplitude $\frac{c_1^2 e}{c_1^2 - \omega^2}$, and the latter an amplitude of $\frac{\omega^2 e}{c_1^2 - \omega^2}$, so that they differ by the amount e ; this shows that OC and CG are in the same straight line when not running at the critical speed.

7. The solutions show firstly that the motion takes place about the position

$$y = -\frac{M\ddot{y}}{\sigma} = -\frac{g}{c_1^2},$$

which we have seen is the statically deflected position of the centre of gravity of a perfectly balanced machine. It is sometimes contended that as the speed increases the rotor shaft tends to straighten out, but there is no indication in the present treatment that this is the case.

8. The free vibration

$$x = N_1\sin(c_1 t - \gamma_1),$$

$$y = N_2\cos(c_1 t - \gamma_1),$$

is the same as for a perfectly balanced rotor and has a

frequency the same as the frequency for stationary vibrations, and, as we shall presently see, if expressed in R.P.M., is the same as the first critical speed. It is, therefore, independent of the speed of rotation of the rotor, and there is no possibility of resonance occurring at $\frac{1}{\sqrt{2}} \times$ the critical speed as suggested by the erroneous method mentioned earlier in this section. The free vibration itself is thus represented by two components having the same frequency, but different amplitudes: it is therefore a central ellipse, the centre being at the point $y = -\frac{Mg}{\sigma}$.

It will be shown later that the free vibration is damped out by friction, so that it has no importance in practice.

9. The forced vibration for the centre C of the shaft is given by

$$x = \frac{\omega^2 e}{c_1^2 - \omega^2} \cos \omega t, \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

$$y = \frac{\omega^2 e}{c_1^2 - \omega^2} \sin \omega t - g/c_1^2. \quad . \quad . \quad . \quad (25)$$

This naturally has the frequency corresponding to the angular velocity of rotation ω, and has a maximum value of $\frac{\omega^2 e}{c_1^2 - \omega^2}$ for each axis.

The motion of the centre is thus a whirl, the radius or amplitude of which is proportional to the out-of-balance and is zero when the rotor is perfectly balanced. A perfectly balanced rotor, therefore, cannot whirl in the manner expressed by equations (24) and (25).

The amplitude of the whirl is also proportional to $\frac{1}{c_1^2 - \omega^2}$ and it thus becomes a maximum when

$$\omega = \pm c_1.$$

(The sign $\pm$ merely indicates that the rotation may be in either direction.)

This value of ω gives the first critical speed, which is thus the same as the stationary frequency for transverse vibrations. Reasons will be given later why the radius of whirl does not become infinite at the critical speed, *i. e.*, why the shaft does not break when the rotor reaches this speed.

10. At the critical speed where $\omega^2 = c_1^2$ equations (20) and (21) become :

$$\ddot{x} + c_1^2 x = c_1^2 e \cos c_1 t,$$

$$\ddot{y} + c_1^2 y = c_1^2 e \sin c_1 t - g,$$

which admit of a solution not involving infinite values, namely,

$$x = t\frac{c_1 e}{2} \sin c_1 t$$

$$= t\frac{c_1 e}{2} \cos\left(c_1 t - \frac{\pi}{2}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

$$y = -t\frac{c_1 e}{2} \cos c_1 t - g/c_1^2$$

$$= t\frac{c_1 e}{2} \sin\left(c_1 t - \frac{\pi}{2}\right) - g/c_1^2. \quad . \quad . \quad . \quad . \quad . \quad (27)$$

Equations (26) and (27) thus give the motion at the first critical speed when friction is ignored. They show that the component along each axis has an amplitude which continually increases in proportion to the duration of the motion, in other words the motion at the critical speed is a spiral of continually increasing radius.

11. From (24) and (25) it will be seen that the phase difference of the motion with respect to ωt changes from zero to 180° as ω passes through the value c_1, *i. e.*, as the speed passes through the critical ; (26) and (27) show that at that speed the motion lags behind ωt by 90°. It will be seen later that when friction is taken into account, the lag increases gradually as the speed increases, being still 90° when $\omega = c_1$.

12. Up to this point we have treated the angular velocity θ as a constant $= \omega$, as on the average it will be in practice. Suppose now that it varies slightly from constancy, so that the angular position ωt becomes $\omega t + u$, where u is so small that its square can be neglected, and we may write $\sin u = u$ and $\cos u = 1$. We then have $\cos\theta = \cos\omega t - u \sin\omega t$ and $\sin\theta = \sin\omega t + u\cos\omega t$, also $\ddot{\theta} = \ddot{u}$. Substituting these values in equations (20), (21), and (15), and writing $\frac{\sigma}{M} = c_1^2$ as before, we get

$$\ddot{x} + c_1^2 x = \omega^2 e \cos\omega t - \omega^2 e u \sin\omega t, \quad . \quad . \quad . \quad (28)$$

$$\ddot{y} + c_1^2 = y\omega^2 e \sin\omega t + \omega^2 e u \cos\omega t - g, \quad . \quad . \quad (29)$$

$$Mk^2\ddot{u} + \sigma e(x \sin\omega t - y\cos\omega t + xu\cos\omega t + yu\sin\omega t) = 0. \quad (30)$$

As a first approximation we substitute in (30) the values already found for x and y, as given in (24) and (25), for the forced vibration and thus obtain (neglecting u in comparison with 1 radian) :

$$Mk_1^2\ddot{u}+\sigma e^2\frac{\omega^2}{c_1^2-\omega^2}u=-\sigma e g/c_1^2\cos\omega t, \quad . \quad . \quad (31)$$

giving for the forced vibration (the free vibration may be ignored as it will be damped out) :

$$u=\frac{e(c_1^2-\omega^2)}{\omega^2e^2c_1^2-\omega^2k_1^2(c_1^2-\omega^2)}\cdot g\cos\omega t. \quad . \quad . \quad (32)$$

At the critical speed $\omega=c_1$, and $u=0$, *i. e.*, the angular velocity of rotation is constant ; at other speeds than those in the neighbourhood of the critical, the term $\omega^2e^2c_1^2$ may be neglected, as e^2 is much smaller than k_1^2, so that

$$u=-\frac{e}{k_1^2}\cdot\frac{g}{\omega^2}\cos\omega t. \quad . \quad . \quad . \quad . \quad . \quad (33)$$

The variation in the angular velocity has thus the same frequency as that of the rotation itself, but is very small in magnitude, as will be seen from the following figures. For a turbo-generator rotor balanced to about 1 oz. at radius k_1 per ton weight of rotor, e/k_1 will be about 3×10^{-5}, and for a machine to run at 3000 R.P.M., k_1 will be about 50 cm., and ω is, say, $2\pi\times50$,

$$u=\frac{3\times10^{-5}}{50}\times\frac{981}{4\pi^2\times50\times50}\cos\omega t$$

$$=6\times10^{-9}\cos\omega t,$$

that is the vibration is very small, the amplitude being only 6×10^{-9} of one radian.

It should be noted that u is proportional to e and to g ; this variation therefore arises through the action of gravity on the rotor when not perfectly balanced, and the variation will be absent if the balance is perfect.

Substituting the value of u in (28) and (29) in order to find the effect of the irregularity on the displacement :—

$$\ddot{x}+c_1^2x=\omega^2e\cos\omega t+\frac{e^2}{k_1^2}g\sin\omega t\cos\omega t,$$

$$\ddot{y}+c_1^2y=\omega^2e\sin\omega t-\frac{e^2}{k_1^2}g\cos^2\omega t-g,$$

or

$$\ddot{x}+c_1^2x=\omega^2e\cos\omega t+\frac{1}{2}\frac{e^2}{k_1^2}g\sin 2\omega t,$$

$$\ddot{y}+c_1^2y=\omega^2e\sin\omega t-\frac{1}{2}\frac{e^2}{k_1^2}g\cos 2\omega t-g\left(1+\frac{1}{2}\frac{e_2}{k_1^2}\right).$$

The solution is (for the forced vibration)

$$x=\frac{\omega^2e\cos\omega t}{c_1^2-\omega^2}+\frac{1}{2}\frac{e^2}{k_1^2}\frac{1}{c_1^2-4\omega^2}\sin 2\omega t, \quad . \quad . \quad . \quad . \quad . \quad (34)$$

$$y=\frac{\omega^2e\sin\omega t}{c_1^2-\omega^2}-\frac{1}{2}\frac{e^2}{k_1^2}\frac{1}{c_1^2-4\omega^2}\cos 2\omega t-\frac{g}{c_1^2}\left(1+\frac{1}{2}\frac{e^2}{k_1^2}\right). \quad (35)$$

The irregularity has thus two effects on the main whirl: firstly, the static deflexion is increased by a small amount, and secondly, there is superimposed on the main whirl a ripple of double frequency, which rises to a maximum at half the critical speed. But the effect is very small, and may not be noticeable; in any case, as the double frequency effect depends on e it cannot appear when the machine is well balanced.

There is, however, the possibility that a rotor which is not perfectly balanced may show vibration at half the critical speed due to the action of gravity, although gravity would produce no such effect at the full critical speed. A vertical spindle rotor is not of course subject to the action of gravity in this sense, and if it vibrated specially at half the critical speed, the cause must be sought elsewhere.

13. We shall now consider some further possible causes of subsidiary critical speeds or speeds where marked vibration may appear other than the normal calculated critical speed.

14. An important case is that of a rotor slotted for a 2-pole winding or with a shaft in which a key-way is cut, where the rigidity of the rotor is greater in one direction than in a direction 90° away, so that if the shaft is rotating, the stiffness in the direction of any one of the axes is not a constant σ but $\sigma+\epsilon\cos 2\omega t$, where ϵ is small in comparison with σ. We shall assume that the rotor is perfectly balanced ($e=0$) and to simplify the examination shall first consider the vertical motion only. The equation is:

$$M\ddot{y}+\sigma y=-Mg-y\epsilon\cos 2\omega t.$$

The first approximation is:

$$y=-\frac{Mg}{\sigma}=-g/c_1^2.$$

Inserting this on the right-hand side of the equation and solving, we get

$$y = -\frac{M_g}{\sigma}\left\{1 + \frac{\epsilon}{\sigma - 4M\omega^2}\cos 2\omega t\right\}$$

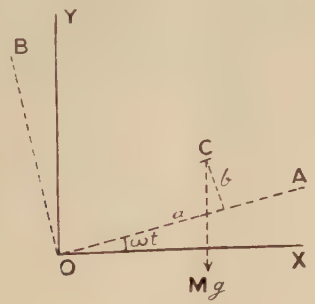
$$= -g/c_1^2\left\{1 + \frac{\epsilon/M}{c_1^2 - 4\omega^2}\cos 2\omega t\right\}. \quad . \quad . \quad (36)$$

There is thus a double frequency vibration about the statically flexed position, which has a maximum when $\omega = \frac{1}{2}c_1$, that is, at half the critical speed.

It is evident that such a motion must have a tendency to arise if a rotor is unsymmetrical as regards its rigidity, for in such a case when the shaft rotates the deflexion will be a maximum or minimum twice every revolution, and if the frequency of the consequent up and down motion is equal to the critical speed there will be resonance; this will be the case whether the rotor is perfectly balanced or not.

It is thus possible for a perfectly balanced rotor which would be quite steady at the critical speed to show marked vibration at half that speed. If the normal running speed is above the critical the forces called into play at half the critical speed will be very small and may give no appreciable effect, but if the running speed is in the neighbourhood of half the critical speed vibration might arise.

Fig. 4.



15. It is worth while to examine the motion a little more fully as there will evidently be some vibration in the horizontal plane also. Let C (fig. 4) be the position of the centre line of the shaft and OA, OB two axes at right angles rotating about O with the same angular velocity ω as rotor. Let the co-ordinates of C be a and b with respect to OA and OB and u and v the corresponding velocities along those axes.

Then $u=a-b\omega$,

$v=\dot{b}+a\omega$,

and the accelerations are

along OA : $\dot{u}-v\omega=a-2\dot{b}\omega-a\omega^2$,

along OB : $\dot{v}+u\omega=\ddot{b}+2\dot{a}\omega-b\omega^2$.

If the force required to produce unit deflexion in the shaft is $\sigma+\epsilon$ along OA and $\sigma-\epsilon$ along OB, and we resolve along OA and OB, the equations are :

$$\mathrm{M}(\ddot{a}-2\dot{b}\omega-a\omega^2)+(\sigma+\epsilon)a=-\mathrm{M}g\sin\omega t,$$

$$\mathrm{M}(\ddot{b}+2\dot{a}\omega-b\omega^2)+(\sigma-\epsilon)b=-\mathrm{M}g\cos\omega t\,;$$

that is,

$$\{\mathrm{M}(\mathrm{D}^2-\omega^2)+(\sigma+\epsilon)\}\,a-2\mathrm{M}\omega\mathrm{D}b=-\mathrm{M}g\sin\omega t,$$

$$\{\mathrm{M}(\mathrm{D}^2-\omega^2)+(\sigma-\epsilon)\}\,b+2\mathrm{M}\omega\mathrm{D}a=-\mathrm{M}g\cos\omega t,$$

giving

$$[\{\mathrm{M}(\mathrm{D}^2-\omega^2)+\sigma\}^2-\epsilon^2+4\mathrm{M}^2\omega^2\mathrm{D}^2]\,a=-\mathrm{M}g(\sigma-\epsilon-4\omega^2)\sin\omega t,$$

$$[\{\mathrm{M}(\mathrm{D}^2-\omega^2)+\sigma\}^2-\epsilon^2+4\mathrm{M}^2\omega^2\mathrm{D}^2]\,b=-\mathrm{M}g(\sigma+\epsilon-4\omega^2)\cos\omega t,$$

The solution is (neglecting ϵ^2 in comparison with σ)

$$a=-\frac{\mathrm{M}g}{\sigma}\left(1-\frac{\epsilon}{\sigma-4\mathrm{M}\omega^2}\right)\sin\omega t,$$

$$b=-\frac{\mathrm{M}g}{\sigma}\left(1+\frac{\epsilon}{\sigma-4\mathrm{M}\omega^2}\right)\cos\omega t.$$

This gives the position with respect to the rotating axes ; the position with respect to the fixed axes is

$$x=a\cos\omega t-b\sin\omega t,$$

$$y=b\cos\omega t+a\sin\omega t,$$

that is :

$$x=+\frac{\mathrm{M}g}{\sigma}\left\{\frac{\epsilon}{\sigma-4\mathrm{M}\omega^2}\right\}\sin 2\omega t, \quad . \quad . \quad . \quad (37)$$

$$y=-\frac{\mathrm{M}g}{\sigma}\left\{1+\frac{\epsilon}{\sigma-4\mathrm{M}\omega^2}\cos 2\omega t\right\}. \quad . \quad . \quad . \quad (38)$$

This result is the same for y as obtained in (34) by the

method of approximations; it shows that there is a similar motion of equal magnitude and 90° out of phase along the horizontal axis, so that the motion is a circular whirl of double frequency, which rises to a maximum at half the critical speed.

16. When the rotor is out of balance the equation for the vertical motion is

$$M\ddot{y} + \sigma y = M\omega^2 e \sin \omega t - Mg - y\epsilon \cos 2\omega t.$$

The first approximation is given by (25) and inserting this on the right-hand side of the above equation we get:

$$M\ddot{y} + \sigma y = M\omega^2 e \sin \omega t - Mg + \frac{Mg}{\sigma}\epsilon \cos 2\omega t + \frac{\epsilon}{2}\frac{M\omega^2 e}{\sigma - M\omega^2}\{\sin \omega t - \sin 3\omega t\}. \quad (39)$$

The first three terms on the right correspond to the main whirl and the double frequency whirl already dealt with. The last term on the right will give in the solution a triple frequency vibration, viz.:

$$-\frac{1}{2}\frac{M\omega^2 e}{\sigma - M\omega^2}\frac{\epsilon}{\sigma - 9M\omega^2}\sin 3\omega t,$$

which has a maximum value at $\frac{1}{3}$ the critical speed. This vibration is, however, proportional to e, the out-of-balance force, and cannot arise in a perfectly balanced machine. The remarks made as to the limited conditions under which the double frequency vibration might arise apply with even greater force to the triple frequency vibration as the damping effect of friction will be correspondingly greater.

17. Another case of interest is that in which covers or sleeves are mounted on the rotor, or the rotor has slots in the periphery for an exciting winding, closed by pressed-in keys; the closeness of these force fits will vary with the deflexion, and the deflexion of the shaft may therefore be not quite proportional to the force applied, *i. e.*, the force to produce a deflexion x will not be σx, but say $\sigma(x + \epsilon x^3)$, where ϵ is small in comparison with unity. (The expression for the force must contain odd powers of x only as the rigidity is symmetrical, that is, the same numerically for the same numerical value of x whether x is positive or negative; if even powers were included this could not be the case as an even power of x is always positive even if x itself is negative.)

The equations then are, putting the small quantities on the right-hand side:

$$\ddot{x} + c_1^2 x = \omega^2 e \cos \omega t - c_1^2 \epsilon x^3,$$

$$\ddot{y} + c_1^2 y = \omega^2 e \sin \omega t - g - c_1^2 \epsilon y^3.$$

Neglecting the small quantities, the forced vibration is as before given by

$$x = \frac{\omega^2 e}{c_1^2 - \omega^2} \cos \omega t, \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

$$y = \frac{\omega^2 e}{c_1^2 - \omega^2} \sin \omega t - g/c_1^2. \quad . \quad . \quad . \quad . \quad (41)$$

Inserting these values on the right-hand side of the original equation, we get after some reduction :—

$$\ddot{x} + c_1^2 x = \{\omega^2 e - \tfrac{3}{4} c_1^2 \epsilon \rho^3\} \cos \omega t - \tfrac{1}{4} c_1^2 \epsilon \rho^3 \cos 3\omega t,$$

$$\ddot{y} + c_1^2 y = -g + c_1^2 \epsilon y_0 (\rho^2 + y_0^2) + \{\omega^2 e + c_1^2 \epsilon \rho (2y_0^2 + \tfrac{3}{4}\rho^2)\} \sin \omega t$$
$$- c_1^2 \epsilon \rho^2 y_0 \cos 2\omega t + \tfrac{1}{4} c_1^2 \epsilon \rho^3 \sin 3\omega t,$$

where $\rho = \dfrac{c_1^2 e}{c_1^2 - \omega^2}$

and $y_0 = g/c_1^2$.

Solving these equations we get:

$$x = \rho \left\{ 1 - \tfrac{3}{4} c_1^2 \epsilon \frac{\rho^2}{c_1^2 - \omega^2} \right\} \cos \omega t - \tfrac{1}{4} c_1^2 \epsilon \frac{\rho^3}{c_1^2 - 9\omega^2} \cos 3\omega t,$$

$$y = -y_0 \{1 - \epsilon y_0 (\rho^2 + y_0^2)\}$$
$$+ \rho \left\{ 1 - \frac{c_1^2 \epsilon}{c_1^2 - \omega^2} (2y_0^2 + \tfrac{3}{4}\rho^2) \right\} \sin \omega t$$
$$- \frac{c_1^2 \epsilon}{c_1^2 - 4\omega^2} \cdot \rho^2 y_0 \cos 2\omega t$$
$$+ \tfrac{1}{4} \cdot \frac{c_1^2 \epsilon}{c_1^2 - 9\omega^2} \rho^3 \sin 3\omega t.$$

Examining these terms in turn we find that the centre of motion is now at the point

$$x = 0,$$
$$y = -y_0 \{1 - \epsilon y_0 (\rho^2 + y_0^2)\},$$

instead of the point $x = 0$, $y = -y_0$. This indicates that the centre of motion rises, *i. e.*, the shaft straightens out slightly, as the vibration increases.

The main vibration, represented by $\cos \omega t$ and $\sin \omega t$, shows a slight change in amplitude, but as before the maximum occurs at the critical speed.

The term in $\cos 2\omega t$ indicates there is a double frequency ripple in the vertical motion y (but not in the horizontal component x) having an amplitude $\frac{{c_1}^2\epsilon}{{c_1}^2 - 4\omega^2}\rho^2 y_0$, that is, $\frac{{c_1}^2\epsilon}{{c_1}^2 - 4\omega^2} \cdot \left(\frac{\omega^2\epsilon}{{c_1}^2 - \omega^2}\right)^2 y_0$. This rises to a maximum at half the first critical speed (when the amplitude changes sign) and again at the first critical speed (when the amplitude does not change sign). Noticeable vibration may thus occur at half the critical speed, but it will take place principally in the vertical plane.

Both components show a triple frequency vibration expressed by $\cos 3\omega t$ and $\sin 3\omega t$, which reaches a maximum at one-third the critical speed, and the amplitude of the vibration changes sign at that point. This vibration also has a maximum value at the critical speed.

Points of marked vibration due to lack of proportionality in the deflexion can thus only show themselves when the machine is out-of-balance, and if they become appreciable at all will only occur at half or one-third, etc., of the critical speed. If, however, these fractions of the critical speed correspond to low running speeds, the forces may be so small as not to produce any noticeable effect.

18. It thus appears that subsidiary critical speeds are only to be expected at half or possibly one-third of the calculated first critical speed, and only then when the subsidiary critical speed is high enough to make the forces appreciable—for example, in the case of a turbo alternator when the speed indicated by the calculation approaches the running speed. All these effects should disappear with perfect balancing, excepting that due to lack of uniformity in the resistance of the shaft or rotor to bending in directions perpendicular to its axis, such as might arise through two-pole slotting of the rotor or through a key-way in the shaft.

The forces tending to produce vibration are small, and the vibrations arise through a kind of resonance; as there is a good deal of damping due to air friction and to the movement of the shaft in the bearing where the oil exercises a strong damping action, the vibrations may not arise at all. This question is gone into more fully in a later section.

19. We have now to consider some cases where resonance may arise from causes outside the machine itself, and two

classes may be noted, firstly where there is an irregularity in the torque applied to the shaft, and secondly where there is resonance with masses outside the machine.

20. Irregularities in the torque driving the machine may arise, for example, through variation in the steam admission or through a fluctuating electrical load.

The result of fluctuation in the torque will be a corresponding fluctuation in the angular velocity of rotation so that the angular position, instead of being ωt, will be $\omega t + \epsilon \sin pt$, where ϵ is a small angle and p is an angular velocity corresponding to the frequency of the disturbance. Then $\cos(\epsilon \sin pt) = 1$ and $\sin(\epsilon \sin pt) = \epsilon \sin pt$. The equations then become, taking the small quantities on to the right-hand side :

$$\ddot{x} + c_1^2 x = \omega^2 e\ (\cos \omega t - \epsilon \sin pt \sin \omega t),$$

$$\ddot{y} + c_1^2 y = \omega^2 e\ (\sin \omega t + \epsilon \sin pt \cos \omega t) - g;$$

that is,

$$\ddot{x} + c_1^2 x = \omega^2 e\{\cos \omega t + \epsilon/2(\cos \overline{\omega + p}\ .\ t - \cos \overline{\omega - p}\ .\ t)\}, \quad (42)$$

$$\ddot{y} + c_1^2 y = \omega^2 e\{\sin \omega t + \epsilon/2(\sin \overline{\omega + p}\ .\ t + \sin \overline{\omega - p}\ .\ t)\}. \quad (43)$$

The main vibration is the same as before, but there are two small vibrations superimposed; the one has a frequency corresponding to $\omega + p$ and a maximum when $\omega = c_1 - p$, the other a frequency corresponding to $\omega - p$ and a maximum when $\omega = c_1 + p$.

This shows that the vibration may have a maximum at speeds corresponding to the sum of and to the difference between the critical speed and the speed corresponding to the frequency of the disturbing fluctuation. So that if disturbed running show itself at such a speed that it cannot be otherwise explained, a cause may be sought for in this direction.

21. The other variety of resonance mentioned is that where, for example, the foundations are not sufficiently rigid and the machine as a whole is vibrating so that there is resonance between the rotor on its shaft and the machine on its foundations. A similar case would be that of a machine rotating in or near a building which itself shows marked vibration, possibly in certain parts only, corresponding to the vibration of the machine. Both these cases are similar in principle and may be illustrated by supposing the whole machine to be mounted on foundations having some elasticity. If then M_1 is the mass of the rotor and M_2 the

effective mass of the machine and that part of the foundations which moves with it, and the forces required to give unity deflexion are in the two cases σ_1 and σ_2 respectively, the equations of motion are as follows:

$$M_1\ddot{y}_1+\sigma_1(y_1-y_2)=M_1\omega^2 e\sin pt,$$

$$M_2\ddot{y}_2+\sigma_2 y_2-\sigma_1(y_1-y_2)=0.$$

giving for the amplitude of the forced vibration

$$y=\frac{\omega^2 e(c_1^2 m+c_2^2-\omega^2)}{\omega^4-\omega^2(c_1^2 m+c_2^2+c_1^2)+c_1^2 c_2^2}, \quad \ldots \quad (44)$$

where $c_1^2=\frac{\sigma_1}{M_1}$, $c_2^2=\frac{\sigma_2}{M_2}$, $m=\frac{M_1}{M_2}$.

Points of marked vibration may thus occur at either of two frequencies given by putting the denominator $=0$; these frequencies will therefore depend on the ratio $\frac{M_1}{M_2}$ as well as on c_1 and c_2, and may thus have almost any values.

For example, if $c_2=c_1$ and $m=0{\cdot}2$, *i. e.*, the mass of the machine and foundations is five times the mass of the rotor—

$$y=\frac{\omega^2 e(1{\cdot}2\,c_1^2-\omega^2)}{\omega^4-\omega_2 c_2^2\times 2{\cdot}2+c_1^4},$$

and a maximum occurs when

$$\omega=c_1\times 1{\cdot}25 \text{ or } c_1\times 0{\cdot}80,$$

that is, at speeds 25 per cent. above and 20 per cent. below the calculated critical speed.

If M_2, the mass of the machine and foundations, is very large in comparison with M_1, the mass of the rotor, the denominator is very nearly equal to $(\omega-c_1)\times(\omega-c_2)$, which shows that in such a case the two speeds where marked vibration may occur nearly correspond to the natural frequency of the rotor and of the machine and foundations respectively. But as the numerator is also small the vibration might not appear if considerable friction is present.

If vibration should occur when $\omega=c_1/\sqrt{2}$, which is, as mentioned above, sometimes thought to be a critical speed, this might indicate that there was resonance with the foundations or some structure outside the machine, in which case,

putting $\omega^2 = c_1^2/2$,

$$c_1^2 - 2\{c_1^2(1+m) + c_2^2\} + 4c_2^2 = 0,$$

that is $$c_2 = \sqrt{\frac{1+2m}{2}}\, c_1,$$

so that if, for example, M_2 were large in comparison with M_1 and m is therefore small,

$$c_2 = c_1/\sqrt{2} = c_1 \times 0{\cdot}707,$$

but if, as in the former example, $M_2 = 5M_1$,

$$c_2 = c_1 \times 0{\cdot}836.$$

B. *Transverse Vibrations with Friction.*

1. The frictional resistance opposing the motion of the rotor may be considered to consist of two parts. The first part opposes the rotation of the rotor about the centre line, C, of the shaft, and this is counteracted by the torque supplied by the turbine, or, if the turbine is cut off from the steam supply, it tends to bring the set to rest; it has no retarding effect on the whirling. The second part opposes the whirling only, and it is with this that we have to deal.

2. It is not known how the frictional resistances opposing the whirling vary with the speed, but it seems likely that they vary with the square of the speed at least. We shall, however, first consider the case where the resistance is assumed proportional to the first power of the speed, as the motion is then simpler to work out, and there is an interesting electrical analogy, which enables the motion to be more readily followed.

The resistance to whirling is in opposition to the path of the rotor centre, so that if $\frac{ds}{dt}$ is the speed of the centre in any direction, the frictional resistance is $M\mu\left(\frac{ds}{dt}\right)^n$, and the components along OX and OY are $M\mu\left(\frac{ds}{dt}\right)^n \frac{dx}{ds}$ and $M\mu\left(\frac{ds}{dt}\right)^n \frac{dy}{ds}$, that is, $M\mu \dot{s}^{n-1}\dot{x}$ and $M\mu \dot{s}^{n-1}\dot{y}$, where μ is a constant.

3. In the particular case we are about to consider, $n=1$, and the components are therefore $M\mu\dot{x}$ and $M\mu\dot{y}$.

The equations of motion are therefore

$$M\ddot{x}+M\mu\dot{x}+\sigma x=M\omega^2 e\cos\omega t,$$
$$M\ddot{y}+M\mu\dot{y}+\sigma y=M\omega^2 e\sin\omega t-Mg.$$

Electrical engineers will notice the similarity between these equations and

$$L\ddot{q}+R\dot{q}+\frac{1}{K}q=E\sin\omega t+E_0,$$

which holds for a circuit comprising an inductance L, a resistance R, a capacity K, an alternating E.M.F. of maximum value E, and a steady E.M.F. E_0, q being the charge in the condenser at any time. Thus, mass is equivalent to inductance, capacity to deflexion per unit force, and applied E.M.F. to applied mechanical force.

The solutions of these equations are, as is well-known :—

$$q=N\epsilon^{-t/T}\sin(pt-\phi)$$
$$+\frac{E}{\sqrt{\{(1/K\omega-L\omega)^2+R^2\}}}\sin\left(\omega t-\tan^{-1}\frac{R}{1/K\omega-L\omega}\right)$$
$$+KE_0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (45)$$

$$y=N\epsilon^{-t/T}\sin(pt-\phi)$$
$$+\frac{M\omega^2 e}{\sqrt{\{(\sigma-M\omega^2)^2+M^2\mu^2\omega^2\}}}\sin\left(\omega t-\tan^{-1}\frac{M\mu\omega}{\sigma-M\omega^2}\right)$$
$$-Mg/\sigma, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (46)$$

that is,

$$y=N\epsilon^{-t/T}\sin(pt-\phi)$$
$$+\frac{\omega^2 e}{\sqrt{(c_1^2-\omega^2)^2+\mu^2\omega^2}}\sin\left(\omega t-\tan^{-1}\frac{\mu\omega}{c_1^2-\omega^2}\right)-g/c_1^2,$$
$$. \quad . \quad . \quad (47)$$

where

$$T=\frac{2L}{R} \quad \text{or} \quad \frac{2M}{M\mu},$$
$$p=\sqrt{1/LK-1/T^2} \quad \text{or} \quad \sqrt{\sigma/M-1/T^2},$$
$$c_1=\frac{1}{\sqrt{LK}} \quad \text{or} \quad \sqrt{\sigma/M}.$$

There is therefore in both cases a free vibration having a frequency slightly less than the natural frequency of the system, but independent of the frequency of the applied E.M.F. or of the speed of the rotor. This vibration is damped out by friction.

The forced vibration is permanent and has a value which is proportional to the applied E.M.F. or out-of-balance force. The vibration is a maximum when $\omega = c_1$, that is the critical speed is the same as the stationary critical, as in the case where friction is ignored ; the amplitude of the vibration at the critical speed is $\frac{E}{R}$ or $\frac{e}{\mu}$, that is, is equal to the applied E.M.F. or out-of-balance force and inversely proportional to the resistance.

The frequency of the forced vibration is the same as the frequency of the E.M.F. or of the rotation, and the charge or displacement lags behind the E.M.F. by an amount depending on the frequency or speed and on the capacity and inductance or elasticity of the shaft and mass of the rotor.

The lag is zero when the frequency is low, but increases to 90° at the critical speed, which, as will be seen, is that corresponding to the natural frequency of the system, while at very high speeds the lag increases to 180°, in other words the force is in opposition to the displacement. The change is similar to that occurring when there is no friction except that in the present case the change is gradual instead of taking place suddenly at the critical speed.

It will be noted that in both cases the vibration takes place about the statically deflected position as a centre. It is evident in both cases that a large static deflexion would increase the tendency to break down, in the one case by puncture or flashing over of the condenser, and in the other by fracture of the shaft.

4. If the resistance to whirling is proportional to the square of the speed, that is $n=2$, the equations are :

$$\ddot{x} + \mu \dot{s} \dot{x} + c_1^2 x = \omega^2 e \cos \omega t,$$

$$\ddot{y} + \mu \dot{s} \dot{y} + c_1^2 y = \omega^2 e \sin \omega t - g.$$

The free vibration (*i. e.* the vibration when $e=0$ or the rotor is perfectly balanced) cannot be expressed in simple terms, but as it will be damped out, as before, it is not of interest.

The forced vibration (*i. e.* the vibration due to the rotor being out-of-balance) is given by

$$x = R \cos(\omega t - \phi), \quad . \quad . \quad . \quad . \quad . \quad . \quad (48)$$

$$y = R \sin(\omega t - \phi) - g/c_1^2, \quad . \quad . \quad . \quad . \quad (49)$$

where R is the radius of whirl of the value :

$$R=\frac{\omega^2 e}{\sqrt{\{\sqrt{(c_1^2-\omega^2)^4+4c^2\omega^4\mu^2\omega^4+(c_1^2-\omega^2)^2\}}} \quad (50)$$

and

$$\tan\phi=\frac{\mu\omega^2 R}{c_1^2-\omega^2}. \quad . \quad . \quad . \quad . \quad . \quad (51)$$

The lag of the displacement behind the force is in this case proportional to the actual deflexion, and in this respect differs from the result obtained in (47). This is of interest as it shows that since the radius of whirling is for a given speed dependent on the out-of-balance, the phase lag will be smaller the more perfectly the machine is balanced ; in the former case where the friction varied as the first power of the speed, the lag was independent of the amount of out-of-balance.

The maximum deflexion occurs, as before, when $\omega=c_1$, that is, when the speed is equal to the stationary critical speed.

5. It is impossible to draw any conclusion from these formulæ as to the real angular advance corresponding to a given speed, as it is not known how the frictional resistance varies with the speed. We can, however, say that if the machine is rotated first in one direction and then in the other, the position corresponding to the out-of-balance will be mid-way between the points of maximum deflexion. When balancing a machine in the running condition it is usual to hold a pencil or chalk against the shaft so that a mark is made on the shaft at a point corresponding to the maximum deflexion. If there were no friction and the speed were not the critical speed, this mark would be in phase with the heavy side of the rotor below the critical speed, and 180° out of phase with it if above the critical speed. But it will be seen from (51) above that the actual position of the mark depends both on the amount of friction and on the amount of out-of-balance. At the critical speed the heavy side of the rotor should be 90° out of phase with the mark on the shaft, but the actual position will be uncertain, as the angle varies rapidly with departure from the critical speed, and it is not usually possible to judge exactly when the machine is running at the critical speed.

6. There is another reason why the position of the mark on the shaft is somewhat uncertain. Referring to fig. 3, if we ignore all other vibrations than that corresponding to the variation in ϕ, we get by taking moments about G :

$$Mk_1^2\ddot{\phi}+\sigma re\sin\phi=0. \quad . \quad . \quad . \quad . \quad (52)$$

If ϕ is a small angle this becomes :

$$Mk_1^2\ddot{\phi}+\sigma re\phi=0,$$

the solution of which indicates a periodic motion having a time of vibration of

$$T=2\pi\sqrt{\frac{Mk_1^2}{\sigma re}}$$

$$=2\pi\sqrt{\frac{k_1^2}{c_1^2 re}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (53)$$

In an actual machine for 3000 R.P.M. we shall have figures of the order of :—$k_1=50$ cm., $k_1/e=3\times10^4$, $c_1=2\pi\times30$, say, while r may be of the order of 1 mm., so that

$$T=\frac{2\pi}{2\pi\times30}\sqrt{\frac{50\times3\times10^4}{0{\cdot}1}}$$

$$=1{\cdot}4\times10^2 \text{ sec.} \quad \text{or about 2 mins.}$$

As the time of vibration is very long in comparison with that of the other vibrations occurring, it will be almost unaffected by the latter, and the assumption that the other vibrations can be ignored, which was made in deducing (52), is therefore justified.

For larger values of e corresponding to less perfect balance and for deflexions of greater magnitude, T will be correspondingly less, and will be greater the more perfect the balance. If friction is ignored, r becomes infinitely great at the critical speed and T becomes zero, and although this can never be the case in practice, it is clear that T may have a value of two mins. or more down to something considerably smaller.

In other words, if the rotor is disturbed from its position of equilibrium by any chance external cause, it may take a considerable time to settle down, and during that period the position of the mark on the shaft will vary considerably from its normal position.

7. At the critical speed the lag is 90°, and the vibration is also a maximum, but the sharpness of this maximum will, as indicated above, depend on the frictional resistance to whirling. In addition to this it will also be influenced by the condition and design of the bearings, as the oil in the bearings exercises a considerable damping influence and also introduces a further complication as follows :

When the speed is low and the vibration therefore small

in magnitude, the film of oil in the bearings will allow the shaft a certain amount of play; this will increase the effective length of the shaft and lower the critical speed; with increasing speed the vibration will therefore start up fairly smartly. As, however, the speed increases and the vibration becomes greater, the shaft may bed hard up against the bearing bush, and increased deflexion will decrease the effective length of the shaft, and so raise the critical speed. As the speed is further increased a similar state of things is gone through, so that at a certain point the vibration will die down more quickly than if there had been no film of oil.

The effect of the oil in the bearings is thus to give an added amount of friction to whirling, and at the same time flatten the maximum peak of the vibration, that is, the vibration will start up and cease fairly smartly and remain more or less constant throughout a fair range of speed. If, however, the film of oil is sufficiently thick or the balance sufficiently good, the vibration may not show itself at all, although it might do so with the same out-of-balance, if the film of oil were thinner.

Section III.—Oscillatory Vibrations—Second Critical Speed.

1. Oscillatory vibration may arise in two ways, either through lack of balance or through vibration transferred from the transverse motion.

2. The lack of balance referred to is of the skew type, that is, is equivalent to a pair of weights at opposite ends of the machine, and 180° apart, giving an out-of-balance couple when the machine rotates; such an out-of-balance will not show itself when the machine is being statically balanced on knife edges, and can only be corrected through observations when the machine is running.

3. Vibration can be transferred from the transverse motion only when the machine is unsymmetrical in the sense that a force applied to the centre of gravity at right angles to the shaft gives a displacement which is not parallel to the centre line, that is, in those cases where, on the static deflexion diagram, the shaft in the deflected position is not parallel to the centre line of the bearings.

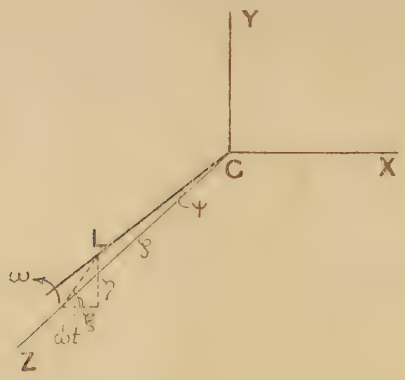
4. The form of out-of-balance mentioned will produce a couple rotating with the machine, that is, a couple alternating with the frequency corresponding to the running speed. Vibration transferred from the transverse motion may, however, be of the frequency corresponding to the

speed of the machine, but may also be of double frequency arising through any of the causes we have discussed. Further, when the normal frequency oscillation has established itself, a double frequency oscillation may start up owing to bipolar asymmetry or some of the other causes mentioned in connexion with the transverse vibration. It is therefore necessary to consider in the oscillatory motion forces both of the actual frequency of rotation and of double frequency.

5. We have found that the transverse motion can, with sufficient accuracy, be considered the same as for a perfectly balanced machine with an out-of-balance force attached to it. In the same way we shall treat the oscillatory motion as being due to an out-of-balance couple of the frequency corresponding to that of rotation or a multiple of that frequency acting on an otherwise perfectly balanced machine.

6. In the diagram fig. 5 let G be the centre of gravity of the rotor and GL the direction of the centre line of the rotor

Fig. 5.



twisted from its normal position by an angle $\psi = \mathrm{LGZ}$, where GZ is the direction of the centre line when not vibrating.

The direction cosines of the centre line GL with the axes GX, GY, and GZ are respectively ξ, η, and ζ: if L is a point at unit distance along the shaft from the centre G, then ξ, η, and ζ are also the co-ordinates of the point L, as shown on the diagram.

If the moments of inertia of the rotor about the shaft centre line and about the line at right angles to it, through

the centre of gravity, are A and B respectively; the angular momenta are *:

$$\text{about GX, } h_1 = \mathrm{B}(\eta\dot{\zeta} - \zeta\dot{\eta}) + \mathrm{A}\omega\xi,$$
$$\text{about GY, } h_2 = \mathrm{B}(\zeta\dot{\xi} - \xi\dot{\zeta}) + \mathrm{A}\omega\eta,$$
$$\text{about GZ, } h_3 = \mathrm{B}(\xi\dot{\eta} - \eta\dot{\xi}) + \mathrm{A}\omega\zeta.$$

In the actual case the angle ψ is very small, so that we can put $\sin\psi = \psi$, $\cos\psi = 1$, and $\zeta = 1$, $\dot{\zeta} = 0$, also the products $\xi\dot{\eta}$ and $\eta\dot{\xi}$ are both negligibly small.

We thus get :

$$h_1 = -\mathrm{B}\dot{\eta} + \mathrm{A}\omega\xi,$$
$$h_2 = \mathrm{B}\dot{\xi} + \mathrm{A}\omega\eta,$$
$$h_3 = \mathrm{A}\omega.$$

* Another and perhaps more legitimate way of deducing these equations is as follows :—

If ω_1, ω_2, and ω_3 are the instantaneous angular velocities about moving axes GX', GY', and GL fixed in the rotor and moving with it : consider the instant when GX is perpendicular to GL and GZ (*cf.* fig. 5), and let θ be the angle between the planes LGZ and YGZ.

Then $\omega_1 = -\dot{\psi}$, $\omega_2 = \dot{\theta}\sin\psi$, and $\omega_3 = \omega$.

The angular momenta about GX', GY', and GZ are $\mathrm{B}\omega_1$, $\mathrm{B}\omega_2$, and $\mathrm{A}\omega$. The angular momenta about the fixed axes GX, GY, and GZ are :

$$h_1 = \mathrm{B}\{\ \omega_1\cos\theta + \omega_2\cos\psi\sin\theta\} + \mathrm{A}\omega_3\sin\psi\sin\theta,$$
$$h_2 = \mathrm{B}\{-\omega_1\sin\theta + \omega_2\cos\psi\cos\theta\} + \mathrm{A}\omega_3\sin\psi\cos\theta,$$
$$h_3 = \mathrm{B}\{-\omega_2\sin\psi\} + \mathrm{A}\omega_3\cos\psi\,;$$

that is

$$h_1 = \mathrm{B}\{-\dot{\psi}\cos\theta + \dot{\theta}\sin\psi\cos\psi\sin\theta\} + \mathrm{A}\omega_3\sin\psi\sin\theta,$$
$$h_2 = \mathrm{B}\{\ \dot{\psi}\sin\theta + \dot{\theta}\sin\psi\cos\psi\cos\theta\} + \mathrm{A}\omega_3\sin\psi\cos\theta,$$
$$h_3 = -\mathrm{B}\dot{\theta}\sin^2\psi + \mathrm{A}\omega_3\cos\psi.$$

Also

$$\xi = \sin\psi\sin\theta \quad \text{and} \quad \dot{\xi} = \dot{\psi}\cos\psi\sin\theta + \dot{\theta}\sin\psi\cos\theta,$$
$$\eta = \sin\psi\cos\theta \quad \text{and} \quad \dot{\eta} = \dot{\psi}\cos\psi\cos\theta - \dot{\theta}\sin\psi\sin\theta,$$
$$\zeta = \cos\psi. \quad \text{and} \quad \dot{\zeta} = -\dot{\psi}\sin\psi\,;$$

so that

$$\eta\dot{\zeta} - \zeta\dot{\eta} = -\dot{\psi}\cos\theta + \dot{\theta}\sin\psi\cos\psi\sin\theta,$$
$$\zeta\dot{\xi} - \xi\dot{\zeta} = \ \dot{\psi}\sin\theta + \dot{\theta}\sin\psi\cos\psi\cos\theta,$$
$$\xi\dot{\eta} - \eta\dot{\xi} = -\dot{\theta}\sin^2\psi.$$

By substituting these values in the equations for h_1, h_2, and h_3, we obtain the relations given above.

The force of restitution of the shaft is :

$$\text{about GX} = \kappa\eta,$$
$$\text{about GY} = -\kappa\zeta,$$
$$\text{about GZ} = 0.$$

If the out-of-balance couple is $\tau\omega^2$, the moment of the couple is :

$$\text{about GX} = -\tau\omega^2 \sin p\omega t,$$
$$\text{about GY} = \tau\omega^2 \cos p\omega t,$$
$$\text{about GZ} = \text{nil},$$

where $p = 1$ or 2.

The equations of motion are thus :

$$-\mathrm{B}\ddot{\eta} - \kappa\eta + \mathrm{A}\omega\dot{\xi} = -\tau\omega^2 \sin p\omega t, \quad . \quad . \quad (54)$$
$$\mathrm{B}\ddot{\xi} + \kappa\xi + \mathrm{A}\omega\dot{\eta} = \tau\omega^2 \cos p\omega t, \quad . \quad . \quad . \quad (55)$$
$$\mathrm{A}\dot{\omega} = 0 ; \quad . \quad . \quad . \quad . \quad . \quad . \quad (56)$$

or putting D for d/dt,

$$(\mathrm{D}^2 + c_2^2)\eta - m\omega\mathrm{D}\xi = \tau\omega^2/\mathrm{B} \,.\, \sin p\omega t,$$
$$(\mathrm{D}^2 + c_2^2)\xi + m\omega\mathrm{D}\eta = \tau\omega^2/\mathrm{B} \,.\, \cos p\omega t,$$
$$\mathrm{A}\dot{\omega} = 0,$$

where $c_2^2 = \dfrac{\kappa}{\mathrm{B}}$ and $m = \dfrac{\mathrm{A}}{\mathrm{B}}$.

The last equation $\mathrm{A}\dot{\omega} = 0$ gives $\omega = \text{const.}$, that is, there is no fluctuation in the angular velocity of rotation.

The other equations give

$$\{(\mathrm{D}^2 + c^2)^2 + m^2\omega^2\mathrm{D}^2\}\,\xi = \tau\omega^2/\mathrm{B}\,\{(c_2^2 - p^2\omega^2) - mp\omega^2\}\cos p\omega t,$$

and a similar equation for ζ in terms of $\sin p\omega t$.

7. The free vibration is not of importance, as it will be damped out as before, but it is of interest to examine its value as illustrating the effect of gyrostatic action on the motion.

The free vibration is of the form

$$x = \mathrm{N}_1 \sin(qt - \phi_1) ; \quad y = \mathrm{N}_2 \cos(qt - \phi_2),$$

where N_1, ϕ_1, N_2, ϕ_2 are constants and

$$\pm q = \tfrac{1}{2}\{m\omega \pm \sqrt{m^2\omega^2 + 4c_2^2}\}. \quad . \quad . \quad . \quad (57)$$

There are thus two natural frequencies of whirling, depending on the direction in which whirling takes place. For example, if the machine is running at what we shall presently see is the critical speed, namely, $m\omega = c_2 = \sqrt{\frac{\kappa}{B'}}$ there are two possible frequencies of whirling for the free vibration, viz.: $q = 1{\cdot}618\ C_2$, or $0{\cdot}618\ C_2$.

8. The forced vibration is given by

$$\xi = \frac{\tau\omega^2/B}{p\omega^2(p-m)-c_2^2}\cos p\omega t$$

$$= \frac{\tau\omega^2}{p\omega^2(pB-A)-Bc_2^2}\cos p\omega t, \quad . \quad . \quad . \quad (57)$$

with a similar equation for η in terms of $\sin p\omega t$.

The amplitude of the vibration is a maximum when

$$p\omega^2(pB-A)-Bc_2^2=0,$$

or

$$\omega^2/c_2^2 = \frac{B}{p(pB-A)} = \frac{1}{p(p-m)} \quad . \quad . \quad . \quad . \quad (58)$$

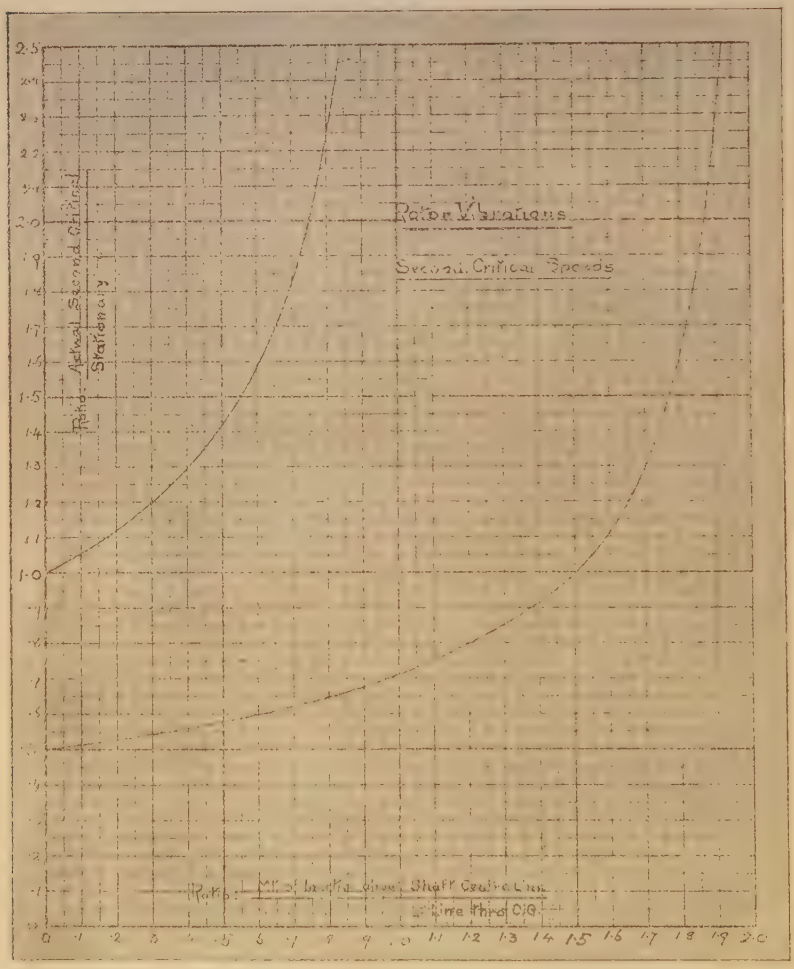
9. This enables the second critical speed to be calculated without difficulty, and fig. 6 gives the necessary curves for reading off the proportional values directly.

The ordinates give the values of ω/c_2 and the abscissæ the values of the ratio A/B. The method of using the curves is as follows:—

1. Work out the radius of gyration k_1 about the shaft centre line.
2. Work out the radius of gyration k_2 about a line perpendicular to this through the centre of gravity.
3. Work out the ratio of A/B or k_1/k_2; for turbo-generator rotors its value is usually between ·2 and ·4 and for flywheels up to 2·0. (This gives the working point on the horizontal axis of the curve.)
4. Work out the "first" critical speed in the usual way and multiply by l/k_2, so as to obtain the stationary "second" critical (where l = half the distance between bearing centres).
5. To obtain the second critical speeds read off from the curves the figures given on the vertical axis and multiply by the stationary critical found from (4).

10. The couple causing oscillatory motion will usually be of a rotating type having components represented by the

Fig. 6.



right-hand side of equations (54) and (55). It is, however, conceivable that one of these components may be absent if the disturbing couple is due to some external cause. Thus if there is a couple round the x axis only, equations (54) and

(55) become

$$-B\ddot{\eta}-\kappa\eta+A\omega\dot{\xi}=-\tau\omega^2\sin p\omega t, \quad . \quad . \quad (59)$$

$$B\ddot{\xi}+\kappa\xi+A\omega\dot{\eta}=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (60)$$

giving the forced vibration

$$\eta=\frac{\tau\omega^2}{B}\frac{c_2^2-p^2\omega^2}{(c_2^2-p^2\omega^2)^2-m^2p^2\omega^4}\sin p\omega t. \quad . \quad . \quad (61)$$

$$\xi=\frac{\tau\omega^2}{B}\frac{mp\omega^2}{(c_2^2-p^2\omega^2)^2-m^2p^2\omega^4}\cos p\omega t, \quad . \quad . \quad (62)$$

The oscillatory motion is thus a whirl the nature of which depends on the speed. The whirl is a maximum when the denominator is zero, that is, at two speeds, one on each side of the stationary critical, and given by

$$\omega^2/c_2^2=\frac{1}{p(p+m)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (63)$$

There is thus a possible further second critical speed, corresponding to the + sign, lower than that already found corresponding to the − sign.

For the ordinary rotating couple the direction of whirling is, of course, always in the direction of rotation, whether the speed is above or below the critical, and this is indicated by the fact that, as will be seen from equations (58), the sign of the amplitude in both planes changes, showing also that the phase of the motion has changed by 180°. But for an alternating couple about the x axis only, as will be seen from (61) and (62), the motion in the horizontal plane changes sign at each of the critical speeds indicated by (63), while the motion in the vertical plane has a further change of sign when $p^2\omega^2=c_2^2$. It will be seen, if these changes are followed out, that the whirling is in one direction below the stationary second critical and in the opposite direction above that speed, while at the stationary second critical the motion is in the horizontal plane only, that is, at right angles to the applied torque.

11. The following example is given to illustrate the application of the above curves and formulæ.

A rotor consists of a solid cylinder 30 ins. diameter and 60 ins. long, running in bearings, 107 ins. between centres. From the deflexion diagram, the deflexion at the centre of gravity is, say, ·0087 in. or ·0221 cm.

We then have:

The first critical speed c_1 is $\frac{300}{\cdot 1485} = 2020$ R.P.M.

The ratio of the radii of gyration k_1/k_2 is $6/19 = 0{\cdot}316$.

Half the distance between bearing centres $= 53{\cdot}5$ ins.

The radius of gyration for the cross moment of inertia $k_2 = 18{\cdot}87$ ins.

The stationary second critical is, therefore,

$$c_2 = c_1 \times \frac{53{\cdot}5}{18{\cdot}87} = 5730 \text{ R.P.M.}$$

From the curves, fig. 6, it will be seen that the main oscillation will occur at

$$5720 \times 1{\cdot}21 = 6940 \text{ R.P.M.}$$

and that a double frequency vibration may possibly show itself at

$$5730 \times 0{\cdot}545 = 3120 \text{ R.P.M.}$$

Thus a turbo-generator designed for a speed of 3000 R.P.M. and having the above mechanical constants might show marked vibration on the overspeed test.

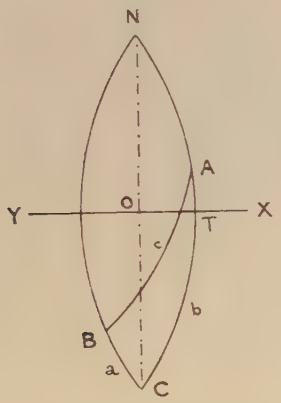
X. *Harmonic Analysis of Motion transmitted by Hooke's Joint.* *By* P. Cormack, *A.R.C.Sc.I., Engineering Dept., Royal College of Science for Ireland* *.

1. With the growth of high-speed machinery, the determination of the accelerations of machine pieces becomes of increasing importance. These determinations are considerably simplified by expressing in the form of a Fourier Series the displacement of the piece under investigation. The value of this method in the analysis of the various phases of the motion of the mechanism of the direct-acting engine is well known. It is here proposed to investigate the coefficients of a Fourier Series for the angular displacement of the driven shaft of a Hooke's Joint. The method being applicable to certain inversions of the slider crank chain, these are also included. The ease with which the coefficients can be determined, and calculations made from the resulting series, make the study of these mechanisms from this aspect one of considerable interest.

* Communicated by the Author.

2. In Hooke's Joint the point B moves in the great circle CBN and the point A in the great circle CAN (fig. 1). The arc AB is constant and equal to a quadrant of the great circle. The point A receives its motion from an arm OA

Fig. 1.



set at right angles to the driving shaft OX, while B transmits motion to the driven shaft OY. Let the angle between the shafts be γ; this will be the angle between the planes of CAN and CBN. In the spherical triangle ABC we have

$$\cos c = \cos a \cos b + \sin a \sin b \cos \gamma.$$

Since $c = \pi/2$, this becomes

$$\cos a \cos b + \sin a \sin b \cos \gamma = 0. \quad . \quad . \quad . \quad (1)$$

Plainly when B is at C, A will be at T; B will therefore move through the angle a while A moves through the angle $b - \pi/2$.

Writing ϕ for a and θ for $b - \pi/2$, equation (1) becomes

$$-\cos \phi \sin \theta + \sin \phi \cos \theta \cos \gamma = 0.$$

Put $\cos \gamma = (1-n)/(1+n)$, and we have

$$(1+n) \cos \phi \sin \theta = (1-n) \sin \phi \cos \theta,$$

$$\therefore \quad n(\cos \phi \sin \theta + \sin \phi \cos \theta) = \sin \phi \cos \theta - \cos \phi \sin \theta.$$

$$\therefore \quad \sin (\phi - \theta) = n \sin (\phi + \theta).$$

$$\therefore \quad e^{i(\phi-\theta)} - e^{-i(\phi-\theta)} = ne^{i(\phi+\theta)} - ne^{-i(\phi+\theta)}.$$

Multiplying both sides by $e^{i(\phi-\theta)}$ gives

$$e^{2i(\phi-\theta)}-1 = ne^{2i\phi}-ne^{-2i\theta}.$$

$$\therefore\quad e^{2i(\phi-\theta)}\{1-ne^{2i\theta}\} = 1-ne^{-2i\theta}.$$

$$\therefore\qquad e^{2i(\phi-\theta)} = (1-ne^{-2i\theta})/(1-ne^{2i\theta}).$$

Since $\phi-\theta$ lies between $+\frac{\pi}{2}$ and $-\frac{\pi}{2}$, and n is less than unity, we may write

$$\begin{aligned}2i(\phi-\theta) &= \log(1-ne^{-2i\theta})-\log(1-ne^{2i\theta})\\ &= -ne^{-2i\theta}-\tfrac{1}{2}n^2e^{-4i\theta}-\tfrac{1}{3}n^3e^{-6i\theta}-\ldots\\ &\qquad +ne^{2i\theta}+\tfrac{1}{2}n^2e^{4i\theta}+\tfrac{1}{3}n^3e^{6i\theta}+\ldots\\ &= n\,.\,2i\sin 2\theta+\tfrac{1}{2}n^2\,.\,2i\sin 4\theta+\tfrac{1}{3}n^3\,.\,2i\sin 6\theta+\ldots\end{aligned}$$

$$\therefore\quad \phi-\theta = n\sin 2\theta+\tfrac{1}{2}n^2\sin 4\theta+\tfrac{1}{3}n^3\sin 6\theta+\ldots\quad\ldots\quad(2)$$

It will be evident that (2) gives the displacement of the driven shaft relatively to the driving shaft. In practice the angle between shafts joined by a Hooke's coupling rarely exceeds 15°. Since $\cos\alpha=\dfrac{1-n}{1+n}$, we have

$$n = (1-\cos\alpha)/(1+\cos\alpha) = \tan^2\frac{\alpha}{2}.$$

For $\alpha=15^\circ$ we get $n=\cdot0173$, so that we can without appreciable error neglect the terms containing the square and higher powers of n in (2) and put

$$\phi-\theta \doteqdot n\sin 2\theta.\quad\ldots\quad(3)$$

For the above value of α, the maximum value of $\phi-\theta$ given by (3) is ·0173 radian or nearly one degree.

From (2) we have

$$\phi = \theta+n\sin 2\theta+\tfrac{1}{2}n^2\sin 4\theta+\tfrac{1}{3}n^3\sin 6\theta+\ldots.\quad(4)$$

$$\therefore\quad \frac{d\phi}{dt} = \frac{d\theta}{dt}(1+2n\cos 2\theta+2n^2\cos 4\theta+2n^3\cos 6\theta+\ldots).\quad(5)$$

$$\therefore\quad \frac{d^2\phi}{dt^2} = \left(\frac{d\theta}{dt}\right)^2(-4n\sin 2\theta-8n^2\sin 4\theta-12n^3\sin 6\theta-\ldots).$$

$$\ldots\quad(6)$$

In obtaining (5) and (6) we assume the series formed by the term-by-term derivative of the member on the right in (4) and (5) to be convergent and to converge to the differential coefficient of the member on the left. In

obtaining (6) we take the case in which the speed of the driving shaft is uniform.

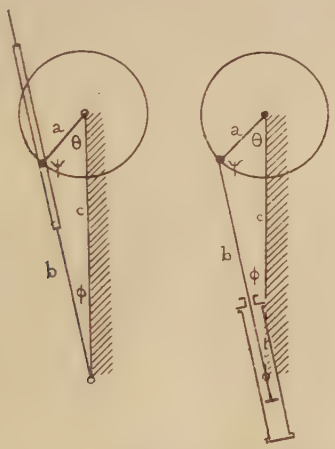
From (5) the maximum speed of the driven shaft is readily seen to be $\omega\{1+2n/(1-n)\}$ or $\omega\sec\alpha$; the minimum speed is seen to be $\omega\{1-2n/(1+n)\}$ or $\omega\cos\alpha$, ω being the angular velocity of the driving shaft.

When the angle between the shafts is not large, (6) may be written

$$\frac{d^2\phi}{dt^2} \doteq -4n\sin 2\theta\left(\frac{d\theta}{dt}\right)^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Thus the maximum angular acceleration of the driven shaft is very approximately $4n\omega^2$. If $\alpha=15^\circ$ and $\omega=60$ radians per sec., the maximum angular acceleration is almost 250 radians per sec. per second.

Fig. 2.



3. In the mechanism of the oscillating cylinder engine, and the quick return crank and slotted-lever mechanism (fig. 2), we have

$$\frac{\sin\phi}{a} = \frac{\sin\psi}{c} = \frac{\sin(\theta+\phi)}{c},$$

or

$$\sin\phi = \frac{a}{c}\sin(\theta+\phi) = n\sin(\theta+\phi), \text{ where } n=\frac{a}{c}.$$

$$\therefore \quad e^{i\phi}-e^{-i\phi} = ne^{i(\theta+\phi)}-ne^{-i(\theta+\phi)}.$$

$$\therefore \quad e^{2i\phi}-1 = ne^{i(\theta+2\phi)}-ne^{-i\theta}.$$

$$\therefore \quad e^{2i\phi}(1-ne^{i\theta}) = 1-ne^{-i\theta}.$$

$$e^{2i\phi} = (1-ne^{-i\theta})/(1-ne^{i\theta}).$$

Since ϕ lies between $+\frac{\pi}{2}$ and $-\frac{\pi}{2}$, and $n<1$,

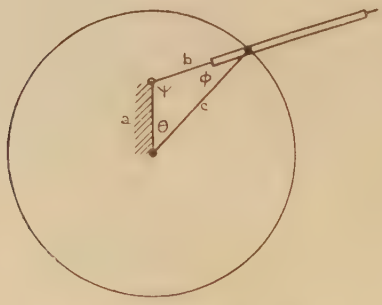
$$\begin{aligned} 2i\phi &= \log(1-ne^{-i\theta}) - \log(1-ne^{i\theta}) \\ &= -ne^{-i\theta} - \tfrac{1}{2}n^2e^{-2i\theta} - \tfrac{1}{3}n^3e^{-3i\theta} - \ldots \\ &\qquad + ne^{i\theta} + \tfrac{1}{2}n^2e^{2i\theta} + \tfrac{1}{3}n^3e^{3i\theta} + \ldots \\ &= n \,.\, 2i\sin\theta + \tfrac{1}{2}n^2 \,.\, 2i\sin 2\theta + \tfrac{1}{3}n^3 \,.\, 2i\sin 3\theta + \ldots. \end{aligned}$$

$$\therefore \quad \phi = n\sin\theta + \tfrac{1}{2}n^2\sin 2\theta + \tfrac{1}{3}n^3\sin 3\theta + \ldots. \quad . \quad . \quad (8)$$

$$\frac{d\phi}{dt} = (n\cos\theta + n^2\cos 2\theta + n^3\cos 3\theta + \ldots)\frac{d\theta}{dt} \quad . \quad . \quad (9)$$

$$\frac{d^2\phi}{dt^2} = (-n\sin\theta - 2n^2\sin 2\theta - 3n^3\sin 3\theta - \ldots)\left(\frac{d\theta}{dt}\right)^2 . \quad . \quad . \quad (10)$$

Fig. 3.



4. In the Pin and Slot mechanism (fig. 3) we have

$$\psi = \pi - \theta - \phi.$$

The angle ϕ is given by (8), so that

$$\psi = \pi - \theta - n\sin\theta - \frac{n^2}{2}\sin 2\theta - \frac{n^3}{3}\sin 3\theta - \ldots, \quad (11)$$

$$\frac{d\psi}{dt} = -\frac{d\theta}{dt}(1 + n\cos\theta + n^2\cos 2\theta + n^3\cos 3\theta + \ldots), \quad (12)$$

$$\frac{d^2\psi}{dt^2} = \left(\frac{d\theta}{dt}\right)^2(n\sin\theta + 2n^2\sin 2\theta + 3n^3\sin 3\theta + \ldots). \quad (13)$$

It will generally be found that we need to consider but the first few terms of these series in making numerical calculations.

XI. *Short Electric Waves obtained by Valves.* *By* E. W. B. Gill, *M.A., B.Sc., Fellow of Merton College, Oxford, and* J. H. Morrell, *M.A., Magdalen College, Oxford* *.

1. THERE have recently been discovered methods for the generation of continuous oscillations of short wave length (of the order of about a metre) by means of three electrode valves. In January 1920, Barkhausen and Kurz † found that with hard valves—*i. e.*, valves at extremely low pressure, if the filament and the plate were approximately the same potential, or, indeed, if the plate were at a potential considerably lower than the filament, provided that the grid was kept at a high potential with regard to them, continuous oscillations could be maintained in a circuit of the Lecher Wire type connected to the grid and plate. The wave length depended primarily on the grid voltage, but also on the emission from the filament and on the plate voltage.

Whiddington ‡ had previously described another method of getting oscillations of lower frequencies using a soft valve, *i. e.*, a valve containing gas at low but appreciable pressure. He employed more usual circuits for a valve, in that the plate was at a high positive potential with regard to the filament and the grid at a few volts above the filament. In this case longer waves were emitted, and he noticed that if V was the grid potential and λ the wave-length emitted, then $\lambda^2 V$ was constant §.

There appear to be other arrangements not hitherto recorded which will also give these waves. With a hard valve and with the grid at a positive potential, oscillations can be obtained if the Lecher Wire system is connected across the filament and grid; the plate may be positive, negative, or at the same potential as the filament, or it may be insulated. Further, the third electrode—the plate—is unnecessary, for oscillations can be sustained by means of a valve consisting of a filament and an anode formed as a spiral of wire concentric with the filament, when these two are connected to the Lecher wires. An intermediate arrangement has been worked successfully in which the wave-length of the diode connected as above is modified by a cylinder concentric with

* Communicated by Prof. J. S. Townsend, F.R.S.

† *Physikalischer Zeitschrift*, Jan. 1920.

‡ Whiddington, 'Radio Review,' Nov. 1919.

§ For a general account of these experiments see 'Radio Review,' June 1920.

the anode, but placed outside the valve and set at various potentials. The best conditions for these cases are still under investigation.

2. Barkhausen and Kurz were apparently unable to give any explanation of the way in which the oscillations were sustained, while Whiddington assumed that the emission of ions from the filament was discontinuous and occurred in bursts. The authors, on the other hand, do not think that any special assumptions are necessary, and that the ordinary conditions for the maintenance of oscillations by continuous emission will account for all the facts they have observed, provided that the time taken by the electrons to pass between the electrodes is taken into consideration, as this time is of the same order as the period of the short waves.

In the present paper only oscillations of the Barkhausen type are considered in detail, but the theory can be extended to cover all the types, and an account of some experiments on the last type (with a diode) will be published later.

It is worth noting that certain writers give the impression that the seat of the oscillations is in the gas or in the electrons in the valve, and that the Lecher wires connected to the valve serve only to demonstrate their existence*. It appears from our experiments that the wires or conductors attached to the electrodes are a necessary part of the oscillatory system. Even with the Lecher wires removed, there will always be some circuit composed of the connecting wires to the batteries or even the valve leads up from the sockets, which will have natural periods of a suitable order for short wave oscillation. This fact seems to have been overlooked in some recent determinations of ionizing potentials, where large emissions from a heated filament were used as a source of electrons. Oscillations will take place even when the valves contain a small amount of gas, but in all the experiments described in this paper gas-free valves were used.

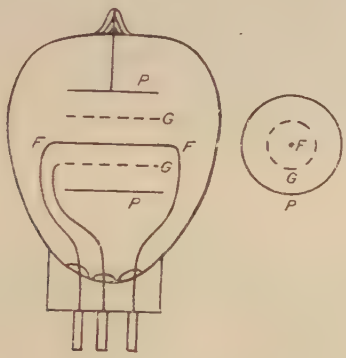
3. It will probably be most convenient first to describe the experiments in detail, and then to set out the theory and apply it to the observed facts.

Various valves were used, but mostly the Marconi M.T.5 valves, which were very kindly given to us by the Marconi Company. These valves consist of a straight filament FF held in the centre of the valve by springy arms. The advantage of the spring is that when the filament is heated and

* Whiddington's theory is independent of there being any external tuned circuit.

expands, the spring prevents sagging. Surrounding the filament is a cylindrical wire grid, GG, composed of thin wire of square mesh, each square having a side of about 1·5 mm. The lead to the grid goes out at the bottom near the filament leads. A cylindrical plate, PP, surrounds the whole with its lead going out through the top of the bulb. These valves being used for transmitting purposes are very thoroughly "glowed out" and pumped to a very high vacuum. The filament emission is very high when heated with 6 volts direct, and for the low emissions that were generally used it was very constant. As the plate lead passes through the top of the bulb, instead of through the bottom and the sealed

Fig. 1.



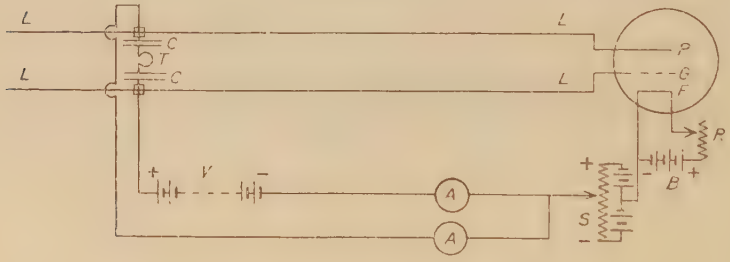
socket, very high insulation is obtained, and, if a strip of tinfoil connected to earth is placed round the outside of the glass, very small anode currents may be measured by an electrometer without any disturbance due to leakage.

It is not necessary for ordinary wireless purposes that the valves should be constructed with the grid and plates either accurately circular in section or accurately centred with regard to the filament; but for the purpose of calculation a symmetrical system of electrodes is necessary and the M.T.5 valve used in most of the experiments was specially selected. All the numerical results to be quoted were obtained from this valve. There is no difficulty in getting the short-wave oscillations with many types of hard valve, the French type produces them quite easily, but the chief reason for selecting the Marconi M.T.5 type was that the electrostatic field between the square-mesh grid and the plate approximates

much more closely to the calculable field between two coaxial cylinders than does the field in the French type, where the grid is a spiral coil of fine wire. The diameter of grid used was 1 cm. and that of the plate was 2·5 cm. to an accuracy of about 5 per cent.

4. The preliminary experiments were made with the apparatus arranged as in fig. 2. The valve is shown diagrammatically: F is the filament, G the grid, P the plate, LL the Lecher wires, which were of copper wire each about 850 cm. long and spaced 5 cm. apart. They were suspended about 200 cm. above the floor from insulators secured to the walls at each end, and from one end were leads about 70 cm. long to the grid and plate respectively. The bridge consisted of two equal condensers, C, C, joined through the heater-coil of a Paul thermo-junction, T. The outer plates were fitted

Fig. 2.



with contacts to slide along the Lecher wires. The capacity of these condensers is unimportant, provided it is large compared with the capacity of the valve. In practice, the capacities were of the order of 1 milli-microfarad. The terminals of the thermo-junction were connected to a galvanometer by two long leads, which are not shown. The sliding contacts were also connected to the negative side of the filament-heating battery B, that on the grid-wire through a high-tension battery V, and that on the plate-wire through a potentiometer S, which could raise the potential of the plate ± 6 volts above the negative end of the filament.

Two sensitive milliammeters, A, A, gave the steady currents through the valve to the grid and filament respectively. A rheostat, R, controlled the filament-heating. In all cases potentials are measured with regard to the negative end of the filament.

With this arrangement the electrons set free at the

filament move outwards under the positive voltage, V, of the grid, and a certain number go direct to the grid and are collected there, the remainder pass through the grid, and, if the potential of the plate is just less than that of the filament, they return to, and are finally collected on the grid. If, on the other hand, the plate potential is a little above that of the filament, a certain proportion of those getting through the grid reach the plate. If the plate potentiometer is now adjusted till the plate current is just zero, and the bridge is moved along the wires, it will be found that with the bridge in certain regions a plate current appears. It was the appearance of this plate current which led Barkhausen to the discovery of the short waves. With the present apparatus these oscillations are also made apparent by the deflexion of the galvanometer attached to the thermo-junction. The positions of the bridge at which the galvanometer gave a maximum deflexion were fairly sharply defined, and did not always coincide with the positions for maximum plate current.

It is not necessary for the plate potential to be so adjusted that the plate current is just zero when oscillations are not occurring. The plate may be set at a considerable negative potential, or the plate voltage may be positive. It was found that for a given grid potential there is a certain plate potential at which the oscillating current through the thermo-junction is a maximum. Also as the potential of the plate was increased, for plate potentials only slightly positive, if oscillations commence the plate current increases; at a certain plate potential no change is noticed in the plate current; and at higher potentials the plate current decreases. For the M.T.5 valve this critical potential was about +2 volts, when the voltage drop down the filament due to the heating current was about 4 volts.

In the first experiments with this apparatus the position of the bridge was varied and the current in the thermo-couple observed when the grid voltage V, the heating current, and the plate potential were all kept constant.

The oscillating circuit consists of a condenser formed by the plate and grid of the valve, the distributed inductance and capacity of the Lecher wires up to the bridge, and the capacities C, C in series with the wires and with the short resistance of the thermal heater which connects them. Hence, if there is an optimum wave-length λ corresponding to the grid voltage V, and if, starting near the valve, the bridge is pushed along the wires, maximum amplitude of oscillation

should occur when the above circuit is tuned to λ, 2λ, 3λ, etc., these positions being indicated by the deflexions of the galvanometer connected to the thermo-junction. Moreover, the distances measured along the wires between successive positions of maximum oscillations should be equal to $\frac{\lambda}{2}$, and all therefore should be equal. It was soon found that this simplicity was not attained, in certain cases equi-spaced positions were found, but in the majority of cases there were at least two sets of positions forming two series of equal spaces, which, as the spacing distance of the two sets was different, appeared to indicate two optimum wave-lengths.

These effects are due to the different modes of oscillation of the system, and, according to the theory which we give below, a grid voltage V will, under suitable conditions, sustain oscillations of short wave-length between certain limits. Any mode of oscillation corresponding to a wave-length between these limits will be maintained. It was therefore desirable to arrange the apparatus so as to avoid these complications.

5. The most obvious improvement was to give up the idea of finding the wave-lengths by moving the bridge, and to put the bridge and its leads at the far end of the parallel wires joined to the valve, and to measure the wave-lengths of the oscillations by means of a loosely coupled secondary circuit. The system of wires connected to the valve is thus fixed. A second pair of long Lecher wires were set up with a loop joining one end, and this loop was brought near the valve circuit. When the secondary is in tune with an oscillation in the primary the current in the primary is reduced. The deflexion of the galvanometer connected to the thermo-junction in the primary circuit may be reduced by 50 per cent. when the bridge in the secondary circuit is in the tuned position, and a movement of 0·5 cm. either way will restore the deflexion to its original value. The distances between the successive positions of the bridge on the secondary circuit, for which the deflexions of the galvanometer attached to the primary circuit are a minimum, are the same, and are equal to half the wave-length of the oscillation in the primary circuit. All the wave-lengths quoted were measured on this form of wave-meter and may be taken as accurate to 0·5 per cent. *

With the condenser bridge and thermo-couple at the far end of the Lecher wires the filament was heated to give an emission of a few milliamperes (this is low heating for an

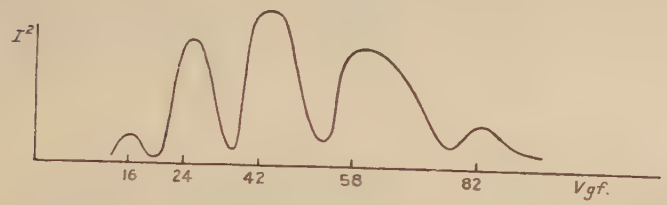
* Townsend and Morrell, Phil. Mag. Aug. 1921, pp. 266–268.

M.T.5 valve) and the grid voltage was raised by two volts at a time by means of batteries of small accumulators from 16 volts to 120 volts, while the plate was kept about 2 volts positive, as this gave large deflexions. The corresponding galvanometer deflexions are shown in fig. 3.

The deflexions are plotted against grid volts ; as a thermal detector was being used, the deflexions are proportional to the mean square of the oscillating current.

The curve shows that oscillations are occurring over nearly the whole range, but that there are maxima for certain voltages—viz., 16, 24, 42, 58, 82, 114, approximately. The wave-lengths measured as above give from 16 to just below 24 volts λ 586 cm., from 24 to 40 volts λ 451, and so on, the wave-lengths for successive portions of the curve being 366, 307, 262, 233. These correspond to the free oscillations of

Fig. 3.



the system, the wires of which were 850 cm. long with leads to the valve about 70 cm. long, with a slight addition for the leads within the valve itself.

The system of wires connected to the valve therefore present a selection of various modes of oscillation with wave-lengths 586, 451, 366, 307, etc., cm., from which the valve chooses the one suitable for the particular voltage V between the grid and plate—the sharp rises just before the various maxima showing that the system oscillates on the longer wave-lengths by preference. For each particular wave-length there is a certain grid voltage which gives the strongest oscillations when the heating current in the filament and the plate voltage is constant ; but the heating current and the potential of the plate relative to the filament both affect the optimum voltage for a given wave-length. Increased emission has the same effect, but this effect depends on the degree of saturation of the emission current.

In the preceding experiments the wave-lengths of the oscillations were measured with a constant heating current

in the filament, but the current from the filament varied with the grid voltage. For the lower voltages all the electrons leaving the filament do not reach the grid space, some returning to the filament. For theoretical reasons it is more convenient to find the grid voltages which give the maximum amplitudes of oscillation on the various wave-lengths when the heating current is so adjusted that the same current flows from the filament to the grid space for all the voltages, the plate voltage being kept constant as before.

The table below gives a set of experiments done under such conditions with an emission current of 6 milliamperes, and the plate at 1·3 volts positive to the filament. In column 1 are given the wave-lengths λ in cms., in column 2 the grid volts V, which excite these wave-lengths most strongly, and in column 3 the product $\lambda^2 V$:—

λ.	V.	$\lambda^2 V$.
208 cm.	156·5	68×10^5
233	122·5	66
262	92·5	64
307	68·5	64·5
366	50·5	67·5
451	36·5	74

All these results, with the exception of the last, agree well with the relation $\lambda^2 V = \text{const.}$

It is not difficult to see why this agreement should be less exact as V decreases. The electrons concerned are not all moving under similar conditions. Owing to the voltage drop of the heating current down the filament, the field between filament and grid differs by about 4 volts for electrons starting from the extreme ends of the filament. And when V becomes comparable to this 4 volts a disturbing factor is introduced.

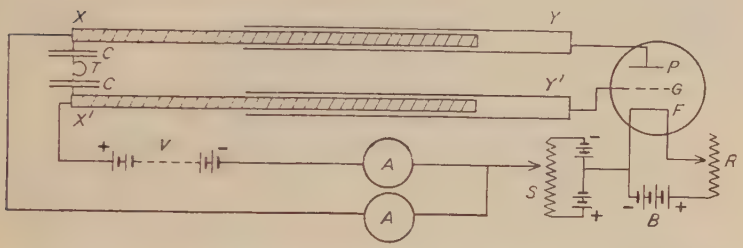
6. These experiments thus give the grid voltages which produce the strongest oscillations on certain definite wave-lengths determined by the particular length of wire used.

To find the range of wave-lengths maintained by a given grid voltage a slightly different apparatus (fig. 4) was used. An adjustable circuit was constructed of two rods, and two telescopic tubes fitted over the rods, so that the effective lengths of the system could be varied by sliding the tubes over the rods.

The condensers and thermo-junctions were attached at the ends X, X^1 of the rods, and the ends Y, Y^1 of the tubes were connected to the plate and grid of the valve respectively, the other connexions being as before.

For brevity, the adjustable circuit will be referred to as the rods. It is not possible to graduate the rods in wave-lengths as against extension of the arms, as this wave-length

Fig. 4.



depends on the emission and on the plate volts. Thus with a fixed length of the arms and 44 volts between grid and plate:

(1) With plate potential fixed.

Emission 2·2 m.a.	$\lambda = 311$ cm.
6·8	306
9·8	300

(2) With emission constant at 5·2 m.a.

Plate potential 1·2 volts.	$\lambda = 308$ cm.
2·4	314

Hence for a given setting λ decreases as the emission rises, and increases as the plate voltage is increased. This is due to the fact that the plate and grid are not a potential node of the oscillating system, but are a variable distance from it depending on the alternating voltage necessary to sustain the oscillation, and this in turn depends on the emission and plate voltage. It is not, however, necessary to go further into this, as the wave-lengths were always found directly by a secondary circuit as in Paragraph 6, the rods being used as a convenient way of varying continuously the wave-length of the system connected to the valve. With all the other factors fixed, the rods were pulled out a centimetre at a time and the oscillating current and wave-length recorded for each position. In one experiment the emission was 1·5 m.a., the grid potential 44 volts, the plate potential 1·8 volts, and oscillations were maintained from $\lambda = 320$ cm. to $\lambda = 451$ cm. with a maximum oscillation about $\lambda = 323$ cm. It was always found that the maximum oscillation was close to the short-wave end of the range.

The effect of (A) varying the emission current keeping the grid and plate voltages constant, and (B) varying the plate voltage keeping the emission current and potential between grid and plate constant was investigated with this apparatus.

In (A) increased emission broadened the range and decreased the wave-length of maximum oscillation.

For example, with $V_{gp}=44$ volts and $V_{pf}=1{\cdot}2$ volts the wave-length for best oscillation with total emission 7·0 m.a. was 295 cm. With total emission 10·6 m.a. it was 274 cm.

In (B) increase of plate voltage increased the wave-length and also broadened the range. Thus with $V_{gp}=44$ volts, and total emission 3·8 m.a. with $V_{pf}=1{\cdot}2$ volts $\lambda=298$ cm., and for $V_{pf}=3{\cdot}0$ volts $\lambda=321$ cm. This last observation must not be confused with the case in which the potential of the plate is increased and that of the grid kept constant. In that case also, increase of plate potential increases the length of the strongest wave, as was observed by Barkhausen, who attributed all the difference in wave-length to the alteration in potential difference between plate and grid. This cannot be the whole of the explanation, for, as stated above, similar results may be obtained by raising both plate and grid equally with respect to the filament.

7. A simple theory to account for the maintenance of the oscillations can be worked out by making some simplifying assumptions; but a general theory will not be attempted, partly because the resistance of the oscillating circuits used was unknown and partly because if the assumptions are not made the calculations become extremely complicated.

These assumptions are :—

A. That the grid and plate can be regarded as forming a parallel plate condenser.

B. That, of the electrons which leave the filament, a fixed small proportion pass through the grid in a uniform stream, and that each electron on passing through the grid has the same velocity.

C. That the electrons which return to the grid from the plate side are nearly all collected directly on it, *i. e.*, only a few pass through on the return journey.

D. That the oscillating potential differences are small compared with the fixed potential differences employed.

It is also assumed that the pressure of the gas inside the valve is so low that the number of collisions between electrons and gas molecules is negligible—this is certainly true for the valves used.

With these assumptions we shall only attempt to show that an oscillation can be maintained of about the right order of wave-length.

The principle involved is the following :—

Suppose the filament and plate are at zero potential and the grid at $+V$; then the electrons from the filament which pass through the grid with a velocity v due to the potential V come to rest at the surface of the plate and return to the grid, which they again reach with velocity v. In the space between the grid and the plate the total work done by the fixed potential V on the electrons which move in this space is zero, all the work having been done between the filament and grid.

If now superposed on the fixed potentials there is an alternating potential $V_0 \sin pt$ between grid and plate due to oscillations, the work done by the potential $V_0 \sin pt$ on the electrons is not necessarily zero. If the work is positive the electrons are abstracting energy from the oscillating system, and the average velocity with which the electrons hit the grid is increased; the oscillations cannot in this case be sustained by the movement of the electrons. But if the work is negative the electrons are giving energy to the oscillating system, and if the rate at which this energy is given is at least equal to the rate of dissipation of energy in the oscillatory circuit by resistance, radiation, or dielectric loss, the valve will maintain the oscillations. The average velocity with which the electrons hit the grid is in this case less than the velocity v due to the potential V which they acquire between filament and grid, and hence the energy put into the system from the battery V is not all used in heating the grid but part is turned into energy of oscillations.

The above argument is not affected if, in consequence of the oscillation, some of the electrons are collected on the plate. In all cases, provided the total work done per oscillation by the alternating field is negative, an oscillation can be sustained if the dissipation of energy in the oscillatory circuit is not large.

8. The particular case in which the filament and the plate are at the same potential when there are no oscillations may be considered first. Let V be the potential above the plate of the grid and d the distance between them.

When there are no oscillations the electron passes the grid with velocity $v\left[=\sqrt{\frac{2e}{m}}\,\mathrm{V}\right]$ and is then subject to a constant retardation f, which brings it to rest just at the plate. If T

is the time the electron takes to pass from grid to plate, $v = f\mathrm{T}$. A further interval of time T brings the electron back to the grid with velocity v.

Assume now that superposed on the fixed potentials is an alternating potential $\mathrm{V}_0 \sin pt$ between plate and grid; the electric force due to this in the space between plate and grid is $\frac{\mathrm{V}_0}{d} \sin pt$, and if $-e$ is the charge on an electron the corresponding force on it is $\frac{e\mathrm{V}_0}{d} \sin pt$ towards the plate.

Since V_0 is taken to be very small compared with V the motion of the electron may to a first approximation be taken as determined solely by V, *i. e.*, its time across is T and retardation f.

The work done by $\mathrm{V}_0 \sin pt$ depends upon the time t_0 at which the electron passes the grid, and for a particular value of t_0 the work is equal to $\int_0^d e \frac{\mathrm{V}_0}{d} \sin pt \, dx$. The axis of x being perpendicular to plate and grid and $x = 0$ being on the grid.

But the velocity at time t is

$$\frac{dx}{dt} = v - f(t - t_0) = f\mathrm{T} - f(t - t_0),$$

and the above work reduces to

$$\int_{t_0}^{t_0+\mathrm{T}} fe \frac{\mathrm{V}_0}{d} (\mathrm{T} + t_0 - t) \sin pt \, dt,$$

which finally gives:

Work on electron going from grid to plate =

$$\frac{2e\mathrm{V}_0}{\mathrm{T}^2} \left(\frac{\mathrm{T} \cos pt_0}{p} + \frac{\sin pt_0 - \sin p\mathrm{T} + t_0}{p^2} \right); \quad . \quad . \quad (1)$$

similarly, the work done on the same electron as it returns from plate to grid comes out as

$$\frac{2e\mathrm{V}_0}{\mathrm{T}^2} \left(\frac{\mathrm{T}}{p} \cos p \overline{t_0 + 2\mathrm{T}} + \frac{\sin p\, t_0 + \mathrm{T} - \sin p\, t_0 + 2\mathrm{T}}{p^2} \right). \quad . \quad (2)$$

Thus the velocities of the electrons on their arrival at the plate or on their return to the grid depends on t_0, that is, on the value of $\mathrm{V}_0 \sin pt$ at the instant they pass through the grid. Assuming a constant stream of electrons through the grid, it is easily seen by integrating (1) for values of t_0 between 0 and $\frac{2\pi}{p}$ that the total work done per period is 0

and similarly for (2). Hence, if all the electrons returned to the grid an oscillation would not be maintained. The possibility of a maintained oscillation depends in this case on the fact that in each oscillation a certain group of the electrons are collected on the plate and the integral of (2) does not in consequence include all the values of t_0 between 0 and $\frac{2\pi}{p}$ and its value is not therefore zero, but may be negative.

The first step is, therefore, to find which electrons reach the plate. When there are no oscillations the electrons have sufficient energy on passing the grid to just take them to the plate against the potential V, and if therefore any extra work is done on them they will be collected on the plate, but if the work is negative they will fall short of the plate and return to the grid. Expression (1) shows that all the electrons which pass through the grid at times t_0, such that $\mathrm{T}\cos pt_0 + \frac{\sin pt_0 - \sin p(\mathrm{T}+t_0)}{p^2}$ is positive, will reach the plate, while those for which it is negative just fail to reach the plate and return to the grid.

Of the electrons then which pass the grid half go on to the plate and half return to the grid, the electrons running to the plate for a time equal to π/p (half the periodic time of the oscillation) and then running back to the grid for time π/p and so on. But the total work done by the oscillating potential on the two halves as they go from grid to plate is zero; and therefore the net work done is the work done on the return journey on the half which returns to the grid.

To find therefore if an oscillation whose periodic time is $2\pi/p$ and amplitude V_0 can be sustained by a grid voltage V it is necessary first to find the time T which the electrons take to pass from the grid to the plate under the field due to V alone, next to find from equation (1) the values of t_0 for those electrons which return to the grid when the system is oscillating, and, finally, by taking the mean value of expression (2) for these values of t_0 and, knowing the emission current, to find the total work done per second by the oscillating potential. If this work is negative and at least equal to the dissipation loss per second, the oscillation will be maintained.

A table of calculated approximate results is given below for various values of the ratio $\mathrm{T} : 1/p$. The second column gives the values of pt_0 for the electrons which return

to the grid in each oscillation, the third gives the total work done per second by the oscillating potential in arbitrary units for a fixed value of V_0 and of emission, and the fourth gives the corresponding wave-lengths for the particular value, V = 44, for which $T = 4{\cdot}3 \times 10^{-9}$ second (see next paragraph):—

	Values of pt_0 for electrons which return to grid.	Work.	V = 44.
$p = \frac{\pi}{4T}$	165° to 345°	negligible.	1040 cm.
$p = \frac{\pi}{2T}$	150° „ 330°	−·47	520 „
$p = \frac{3\pi}{4T}$	135° „ 315°	−·85	347 „
$p = \frac{\pi}{T}$	120° „ 300°	−·36	260 „
$p = \frac{2\pi}{T}$	90° „ 180°	−·32	130 „

In all these cases the work is negative and oscillations can be theoretically sustained, though in practice the dissipation losses in the oscillatory circuits are such that generally only those wave-lengths corresponding to the larger values of the work exist.

For all values of the ratio $T : 1/p$ outside the range of the table the work is small. The larger the work the greater the amplitude of oscillation that will be sustained, and the periodic time $2\pi/p$ of the oscillation of maximum amplitude for a given value of T is seen to be in the region of $\frac{8}{3}$T. In general, the wave-length corresponding to this will be sustained, and also a certain range of wave-lengths on both sides of it, the limits of the range being determined by the dissipation losses in the oscillatory circuit. The theoretical result is in good agreement with several of the experimental results of paragraph 6.

In the particular case recorded there for V = 44 the range of wave-lengths sustained was from 320 to 451 cm., with a maximum amplitude for 323 cm. Increase of emission broadened the range of wave-lengths sustained, which is in accordance with the fact that for a given V_0 the work put into the oscillatory system is proportional to the emission current.

The fact that the wave-length of maximum amplitude of oscillation was near to the short wave-length end of the range also agrees with the calculated fact that the

work done falls off much more rapidly on the short-wave side of the maximum than on the long-wave side.

9. To calculate the time T an electron from the filament takes to go from the grid to the plate when the grid potential is V volts above both filament and plate it is not necessary to assume the grid and plate to be parallel, but they may be taken, as they actually are, to be concentric cylinders of radii a, b.

The retarding force on the electron when it is at distance v from the axis is $\frac{k}{r}$, where $k = \mathrm{V}/\log_e \frac{b}{a}$.

The equation of motion is therefore

$$m\frac{d^2r}{dt^2} = -\frac{ek}{r},$$

which gives when integrated twice, remembering that $\frac{dr}{dt} = 0$, when $r = b$,

$$\mathrm{T} = b\sqrt{\frac{2m}{e\mathrm{V}}\log_e\frac{b}{a}}\int_0^{\sqrt{\log_e \frac{b}{a}}} e^{-x^2}dx.$$

In the actual valve used $a = \cdot 5$ cm., $b = 1\cdot 25$ cm., and, taking $\frac{e}{m} = 5\cdot 3 \times 10^{17}$ E.S. units and measuring V in volts,

$$\mathrm{T} = \frac{2\cdot 87 \times 10^{-8}}{\sqrt{\mathrm{V}}} \text{ second},$$

the accuracy of this being limited by the accuracy to which a and b are known and probably from 5 to 10 per cent.

The wave-length for any relation between T and $1/p$ can now be at once calculated. If $p = n\pi/\mathrm{T}$, the time of one oscillation is $2\pi/p$ or $2\mathrm{T}/n$ and the wave-length in cm. is $6 \times 10^{10}\mathrm{T}/n$.

The simple theory shows that for the oscillations of maximum amplitude $p\mathrm{T}$ has a certain value about $3\pi/4$. But $\mathrm{T} \propto \frac{1}{\sqrt{\mathrm{V}}}$ and $\lambda \propto \frac{1}{p}$, and hence the connexion between the grid voltage V and the wave-length λ of maximum oscillation is $\lambda^2\mathrm{V} = \text{constant}$.

10. The theory is thus in good general agreement with the experimental results, but there is one fact unaccounted for—that being the variation in the wave-length of the

oscillation of maximum amplitude, for a fixed potential between grid and plate, when either the emission is altered or the plate potential is slightly altered with respect to the filament.

There is also a special case, which is forming the subject of a separate investigation, in which, when the plate is very negative (40 volts or so) with respect to the filament, oscillations can still be produced, but without any current reaching the plate at all. These oscillations are, however, very much weaker and more difficult to produce than those dealt with in our experiments.

The simple theory which depends on the collection on the plate in each oscillation of a group of electrons will obviously not account for this special case.

The explanation of the above considerations is to be looked for in the assumptions made in the simple theory. The first assumption that the grid and plate could be regarded as parallel is not important, as the field between cylinders of the size of the grid and plate used is not far from uniform. (It will be noted that the value of T was calculated for the valve used by taking the field between cylinders.) The only difference between cylinders and parallel plates on the simple theory would be to make the ratio T to $1/p$ for maximum oscillation slightly different. But the second assumption that the electrons pass the grid in a constant stream all having the same velocity requires more careful examination. This velocity is not actually the same for two reasons :

A. Because there will be alternating potentials between the filament and the grid which set up a velocity distribution at the grid.

B. Because of the voltage drop of the heating current down the filament.

In the usual methods of producing oscillations by means of valves the alternating potential of (A) is most important, as it controls the whole action of the valve, but in our experiments it is only of secondary importance. The alternating potentials induced between filament and grid are smaller than those between grid and plate, and the major part of the work done by the alternating field on the electrons, which is what determines whether the electrons reach the plate or not, is done between grid and plate, and it is therefore nearly correct to say that all the electrons passing the grid at times t_0, such that expression (1) is positive, reach the plate.

In the extreme case, however, when the plate is so

negative that the system is unable to oscillate by the method of driving groups of electrons on to the plate, the oscillation is almost certainly due to a velocity distribution at the grid, as this means that the electrons do not pass the grid in a uniform stream, and allows the integral of expression (2) to be finite and not zero, as it normally is when all the electrons return to the grid.

The comparative weakness of the oscillations in this case shows that the electrons have all nearly the same velocity when passing the grid.

The simple theory should therefore be in agreement with the observed facts, as it is when the oscillations are mainly due to the collection by the plate, but as the plate is made more negative with respect to the filament the velocity distribution at the grid becomes more important and the simple theory is less accurate.

The velocity distribution at the grid will also be affected by the emission, as this varies the space charge round the filament—this affecting the time the electrons take to pass from the filament to grid,—and this in turn varies the small effect of the alternating field in this space.

The effect of the voltage drop of about 4 volts down the filament is that, instead of dealing with one stream in the field due to the grid being charged to V volts, there are a series of streams moving under potentials varying from V to V−4 (V being the potential difference between the grid and the negative end of the filament). The number of electrons in the various streams varies from a maximum number corresponding to V−2, the middle of the filament being the hottest. The emission falls off equally on both sides of this middle point.

In the general case, when the plate is slightly positive with regard to the negative end of the filament when there are no oscillations, some of the streams reach the plate and the remainder approach it closely, but to varying distances.

If oscillations commence some of these latter streams are periodically diverted to the plate, while in the other half oscillations some of the former are diverted off.

Thus all the streams concerned maintain the oscillation as in the simple theory, and unless V is small the wave-lengths they each maintain best are nearly the same, so that the combined effect differs little from that of a single stream moving under potential V.

The question of whether the mean plate current rises or falls when oscillations begin depends on whether the average

density of the streams diverted to the plate exceeds or is less than that of those diverted from it.

If the plate is at the same potential as the centre of the filament, no change should therefore occur. If it is above this the current should drop, and if below the current should rise, which is in agreement with the results in paragraph 4.

In conclusion, we should like to express our thanks to Professor Townsend, who has assisted us with much valuable advice and criticism.

Note on the Determination of Ionizing Potentials.

In the experiments above described, the electric fields in the spaces between the grid, filament, and plate are similar to those used in experiments on the determination of critical potentials when a small quantity of gas is introduced.

In the latter experiments variations in the plate current are observed as the grid potential is raised, and at certain potentials of the grid abnormal variations in the plate current are observed which are interpreted as indicating certain critical potentials, characteristic of the molecules of the gas.

The experiments which are here described show that abnormal variations in the plate current are to be expected, due to oscillations which may be maintained where large currents of the order of a milliampere flow from the filament towards the grid when the gas is at a very low pressure. The effect of varying the potential of the grid is clearly shown by the curve of fig. 3.

In all the ordinary methods of wiring the valve to the cells and galvanometers or electrometers, the system seems to be as often in a state of oscillation as not.

Even if the plate is so negative that the oscillations are of the weaker type which do not affect the plate current, the difficulty arises that the oscillations superimpose an unknown potential difference between filament and grid, and the electrons are not moving under the fixed field alone.

It is necessary therefore, in order to obtain reliable values of the critical potential, to take precautions to prevent oscillation, which is best done by using emission currents much less than a milliampere.

XII. *Pseudo-Regular Precession.*

By Sir G. Greenhill *.

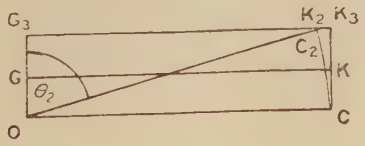
THIS is the gyroscopic motion described and illustrated in Klein-Sommerfeld's *Kreisel-Theorie*, p. 209, where a spinning top appears at first sight to be moving steadily in uniform precession at a constant angle with the vertical, but on closer scrutiny the axle is seen to be describing a crinkled curve of small loops or waves; so that in this penultimate state a realisation is obtained of a motion expressible by a function which does not possess a differential coefficient, a paradox fascinating to a certain school of pure mathematicians.

A pseudo-regular precession, although invisible, would not feel impalpable to the analytical thumb passed over it, which would detect a roughness. But in regular precession the curve would feel quite smooth.

In the investigation the axle may first be supposed moving in perfect steadiness with no trepidation or nutation; and then to receive a small impulse, blow or couple, giving rise to the pseudo-regular precession visible to the eye.

1. Begin with a rapid spin about the axle, held fixed at a constant inclination, taken at first as horizontal for simplicity, in fig. 1.

Fig. 1.



For visible experimental illustration, it is convenient to take a large (52-inch) bicycle wheel, mounted on a spindle with ball bearings, and to prolong the spindle by screwing on a stalk, at one end or both.

The end of the stalk may be supported on the hand and the wheel set in rotation by a swirl; the hand accompanies the wheel in the precession; or else the point may be placed in a cup fixed on the floor.

When the axle is released from rest, it will start from a cusp and sink down, then rise up again in a regular series of loops or festoons; so that to secure the uniform precession

* Communicated by the Author.

an impulse couple must be applied, given by a horizontal tap of appropriate amount.

The word moment or momentum is of such frequent occurrence in dynamical theory that we prefer to replace it by *impulse* instead of momentum, linear or rotational.

Representing then the impulse (rotational) CR due to the rotation R about the axle by the vector OC, the impulse required to start the rotation from rest, or reversed to stop it again, the axle OC in steady motion will move round the vertical OG at a constant (horizontal) inclination, with precession μ, such that the vector velocity of K, the end of the resultant impulse OK, is equal to the impressed couple of gravity; then $CR\mu = gMh$, Mh denoting the preponderance or first moment about O.

This result, true accurately when the axle OC is horizontal (fig. 1), is obtained at any other inclination θ of the axle with the vertical (nadir or zenith, figs. 2, 3) in the elementary Kindergarten treatment, where the top is supposed spinning so fast that the deviation is insensible of axial impulse OC and resultant impulse OK, and then the velocity of C may be equated to the gravity couple, making $CR\mu \sin\theta = gMh \sin\theta$, as before, when the axle was horizontal, on dividing out $\sin\theta$.

2. Hanging down inert, vertically from O in fig. 2, the top forms a compound pendulum of S.E.P.L. (simple-equivalent-pendulum-length) $OP = l = A/Mh$, A denoting the transverse-diametral M.I. (moment of inertia) at O, and C as above the axial M.I.; and in small invisible oscillation in a plane, the top will swing as a pendulum, and beat n/π times a second, where $n^2 = g/l$, $An^2 = gMh$, or make a swing or beat in π/n seconds; Mh may be called the preponderance about O.

Falling down from rest from the upward vertical position, the top will have acquired the angular velocity ω in the lowest position, such that by the Energy-Principle,

$$\tfrac{1}{2} A\omega^2 = 2gMh,$$

and the equivalent rotational impulse $A\omega = 2An$, while $An^2 = gMh$ is the equivalent of P in the *Kreisel-Theorie*.

The impulse $2An = 2\sqrt{(gMhA)}$ is a dynamical constant of the top, and to a geometrical scale may be represented by a length k, in addition to the S.E.P.L. $OP = l$ measured along OC; and then if in any assigned top motion the constant impulse component G and CR, about the vertical OG and

the axle OC, is represented to the same scale by ∂ and ∂',

$$\frac{\partial, \partial'}{k} = \frac{G, CR}{2An} = \frac{h, h'}{n}$$

in Darboux's notation (a different use of h from that employed above).

Time can be reckoned in the pendulum beat, π/n seconds; and the relation, $CR\mu = gMh = An^2$, can be written

$$\frac{\mu}{n} = \frac{A}{C}\frac{n}{R};$$

or expressed in words, the number of beats per circuit of the axle is C/A times the number of revolutions of the top per double beat.

The resultant impulse vector being OK, the component perpendicular to the axle, if horizontal, as in fig. 1, is

$$A\mu = \frac{gMhA}{CR} = \frac{A^2n^2}{CR},$$

or to the geometrical scale, $OC \cdot CK = \frac{1}{4}k^2$, in the steady, regular precession.

For brevity we are allowed to assume tacitly the geometrical scale, and to replace any dynamical quantity in an equation by its vector length, such as the axial impulse CR by the vector length OC, or ∂'.

3. To change this steady motion of the axle into a penultimate pseudo-regular precession, another impulse is applied about a vertical axis, supplied by a horizontal tap on the axle perpendicular to the plane OCK, in fig. 1.

This will cause CK to grow to CK_3, and the resultant impulse to change from OK to OK_3; and to make the pseudo-regular precession advance through a series of cusps, we find that $KK_3 = CK$, and the axle rises from OC to OC_2 at an angle θ_2 with the upward vertical, zenith; where C_2 reaches the level GK_3 of K_3; and here θ_3, the inclination in the lowest position, is $\frac{1}{2}\pi$.

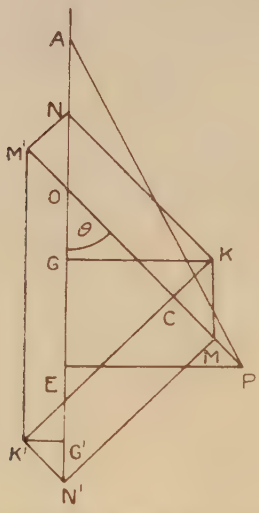
By a general dynamical principle

$$OK_3^2 - OK_2^2 = \tfrac{1}{2}k^2(\cos\theta_2 - \cos\theta_3)$$
$$= 2OC \cdot CK(\cos\theta_2 - \cos\theta_3).$$

For in the general unsteady motion of the axle of a top, where the inclination θ is varying, a new component KH is added to the impulse OK perpendicular to the vertical plane

GOC, of magnitude $A\frac{d\theta}{dt}$, and then the resultant impulse vector OH describes a curve in a horizontal plane GHK, with velocity equal to the gravity couple $g\mathrm{M}h\sin\theta$.

Fig. 2.



The radial velocity of H in the horizontal plane is then

$$\frac{d\mathrm{GH}}{dt} = g\mathrm{M}h\sin\theta\cos\mathrm{GHK} = g\mathrm{M}h\sin\theta\frac{\mathrm{HK}}{\mathrm{GH}},$$

$$\mathrm{GH}\frac{d\,\mathrm{GH}}{dt} = g\mathrm{M}h\mathrm{A}\sin\theta\frac{d\theta}{dt},$$

and integrating,

$$\tfrac{1}{2}\mathrm{GH}^2 = g\mathrm{M}h\mathrm{A}(\mathrm{E}-\cos\theta) = \mathrm{A}^2n^2(\mathrm{E}-\cos\theta)$$
$$= \tfrac{1}{4}k^2(\mathrm{E}-\cos\theta), \quad \tfrac{1}{2}\mathrm{OH}^2 = \tfrac{1}{4}k^2(\mathrm{F}-\cos\theta),$$

to a geometrical scale, with E, F dynamical constants.

The component HK is zero in the upper and lower position, where $\theta=\theta_2, \theta_3$.

In the general unsteady motion of the top, the impulse vector moves from OK to OH with $\mathrm{KH}=\mathrm{A}\frac{d\theta}{dt}$; and with

$$\mathrm{KH}^2 = \mathrm{OH}^2 - \mathrm{OK}^2 = \mathrm{OH}^2 - \frac{\mathrm{GC}^2}{\sin^2\theta},$$

$$\left(\mathrm{A}\frac{d\theta}{dt}\right)^2 = 2\mathrm{A}^2n^2(\mathrm{F}-\cos\theta) - \frac{\mathrm{OG}^2 - 2\,\mathrm{OG}\,.\,\mathrm{OC}\cos\theta + \mathrm{OC}^2}{\sin^2\theta},$$

or with $\cos\theta = z$, and replacing OG, OC for dynamical homogeneity by Darboux's $2Ah$, $2Ah'$,

$$\left(\frac{dz}{dt}\right)^2 = 2n^2(F-z)(1-z^2) - 4(h^2 - 2hh'\cos\theta + h'^2) = 2n^2 Z,$$

thus defining z as an elliptic function of t.

Resolved into factors, we write

$$Z = z_1 - z \,.\, z_2 - z \,.\, z - z_3,$$

in the sequence $\infty > z_1 > 1 > z_2 > z > z_3 > -1$; and then

$$z = z_2 \,\mathrm{sn}^2 \tfrac{1}{2} mt + z_3 \,\mathrm{cn}^2 \tfrac{1}{2} mt, \qquad m^2 = 2n^2(z_1 - z_3).$$

4. Here with OC horizontal in fig. 1, $\cos\theta_3 = 0$,

$$OK_3^2 - OC_2^2 = G_3K_3^2 - G_3K_2^2 = CK_3^2$$
$$= 2OC \,.\, CK\cos\theta_2 = 2CK_3 \,.\, CK, \qquad CK_3 = 2CK.$$

A greater impulse would make the cusps open out into loops in the pseudo-regular precession; but the cusps would be blunted into waves if the impulse was reduced.

Reverse this tap, and K is brought back again to C, and the axle would fall as at first from a cusp and rise again.

In the first cusp motion where the axle rises to a series of cusps and sinks again to the horizontal, the motion is found to be pseudo-elliptic and can be expressed in a finite form,

$$\sin\theta \exp(\psi - ht)\, i$$
$$= \surd(1 - \cos\theta_2 \cos\theta) + i\surd(\cos\theta_2 \cos\theta - \cos^2\theta),$$

connecting azimuth ψ with θ, the inclination to the zenith. The verification is left as an exercise. Here $h = h'\cos\theta_2$.

In the second cusp motion, where the axle is horizontal and falls from a cusp, and then sinks down to an angle θ_3 with the downward vertical, nadir, the (ψ, θ) motion is not pseudo-elliptic; but azimuth ψ and hour angle ϕ change place (θ, ϕ, ψ the Eulerian angles), and

$$\sin\theta \exp(\phi - h't)\, i = \surd(\sec\theta_3 - \cos\theta_3 \,.\, \cos\theta)$$
$$+ i\surd(\cos\theta_3 - \cos\theta \,.\, \cos\theta + \sec\theta_3),$$

where Darboux's $h' = \dfrac{CR}{2An} = \surd(\tfrac{1}{2} \,.\, \sec\theta_3 - \cos\theta_3)$ changes place with h, or ∂' with ∂ ; $h = h'\cos\theta_2$ and ∂ now zero.

But an interchange again of ϕ, ψ will give the (ψ, θ) motion of a non-spinning gyroscopic wheel, or spherical

pendulum, with h', CR and ∂' zero, in which the axle is projected horizontally, with angular velocity $2h=\sin^2\theta\frac{d\psi}{dt}$, and sinks down to an angle θ_3 with the nadir, rising up again to the horizontal, and this makes

$$h=n\surd(\tfrac{1}{2}\,.\,\sec\theta_3-\cos\theta_3).$$

The motion can be shown with a plummet on a thread, say about 10 inches long, to beat as a pendulum twice a second, a double beat period of one second ; whirled round swiftly, the thread rising to the horizontal position, and sinking down again periodically.

Then we find

$$v_2^2=4l^2h^2=2gl(\sec\theta_3-\cos\theta_3),\qquad v_3^2=2gl\sec\theta_3,$$

and in the conical pendulum, at angle θ_2, $v^2=\frac{1}{2}v_2^2$.

The apsidal angle is found to be

$$\Psi=\tfrac{1}{2}\pi+\mathrm{K}\surd(1-2\kappa^2)\rightarrow\pi(1-\tfrac{3}{8}\cos^2\theta_3)$$

as the plummet is whirled round faster.

5. But next suppose the axle OC is held at an angle θ with the zenith, the wheel spun with impulse CR=OC, and then released, in fig. 3.

The axle will start from a cusp, at $\theta=\theta_2$, and the motion in general is not expressible in finite terms as pseudo-elliptic ; but it will represent a gravity brachistochrone on a sphere.

To make the axle move steadily at the inclination θ with constant precession μ, the impulse vector CK is applied perpendicular to the axle, such that $\mu\sin\theta$ being the component rotation of the wheel about the axis OK′ perpendicular to OC, $\mathrm{CK}=\mathrm{A}\mu\sin\theta$ (the inertia of the stalk being ignored), $\mathrm{MC}=\mathrm{A}\mu\cos\theta$, $\mathrm{MK}=\mathrm{A}\mu$, KM drawn vertical to meet OC in M, with the condition

$$\mathrm{GK}\,.\,\mu=\text{gravity couple}=g\mathrm{M}h\sin\theta=\mathrm{A}n^2\sin\theta,$$

$$\mathrm{OM}=\frac{\mathrm{GK}}{\sin\theta}=\frac{\mathrm{A}n^2}{\mu};$$

then the geometrical relation $\mathrm{CM}+\mathrm{MO}=\mathrm{OC}$ becomes

$$\mathrm{A}\mu\cos\theta+\frac{\mathrm{A}n^2}{\mu}=\mathrm{CR},$$

the condition for steady motion. Also

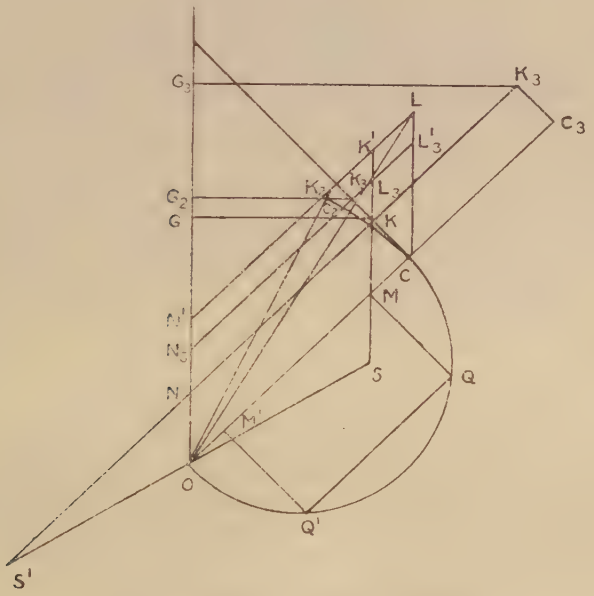
$$\mathrm{OM}\,.\,\mathrm{MC}=\mathrm{A}^2n^2\cos\theta=\tfrac{1}{4}k^2\cos\theta,\quad \mathrm{OM}\,.\,\mathrm{MK}=\tfrac{1}{4}k^2,$$

so that K lies on this hyperbola with asymptotes OC, OG.

Or the position of M is determined by drawing QQ′ parallel to OC, to cut the circle on the diameter OC in Q, Q′, when $MQ^2 = \frac{1}{4}k^2 \cos\theta$.

This is for the small value of μ in fig. 3, and the quiet precession when M is taken close to C, the other point Q′ will determine a motion where the precession μ is swift and the motion violent.

Fig. 3.



$OC_3 = 2OM.$ $MK' = 2MK.$

Spinning upright with $\cos\theta = 1$, $\mu + \frac{n^2}{\mu} = \frac{C}{A}R$; and the two values of μ will give the independent normal invisible circling of the axle round the vertical; there will be $\frac{n^2}{\mu^2}$ rapid beats for one slow beat of the axle.

The general slight oscillation will be compounded of these two circlings, adding up to an epicyclic motion of the axle, a result obtained in this manner without any appeal to approximation.

In fig. 2, the axle is pointing downward at an angle θ with the nadir; here the slow precession of K is retrograde, but in the swift violent motion of the associated K′, the precession is direct.

6. Next to make the axle rise to a cusp on $\theta=\theta_2$ from $\theta=\theta_3$ in the penultimate pseudo-regular precession, the impulse applied is KK_3, to make the axle come to rest on the horizontal G_2K_3 in K_2, C_2, in fig. 3.

Then in the general formula, or on the figure, with CD the perpendicular on K_3M_3,

$$G_2K_3^2-G_2K_2^2=OK_3^2-OK_2^2=CK_3^2=\tfrac{1}{2}k^2(\cos\theta_2-\cos\theta_3),$$

or
$$CK_3^2=K_3D\,.\,K_3M_3=2OM\,.\,MK\,\frac{K_3D}{OC},$$

and producing MK double length to K′,

$$K_3M_3\,.\,OC=2OM\,.\,MK=OM\,.\,MK',$$

implying that if OCLN′ is the parallelogram on OC, the diagonal OL will cut MK′ in L_3 such that $ML_3=M_3K_3$, and K_3 is determined by drawing L_3K_3 parallel to OC, cutting off the length CK_3 on CK, in fig. 3.

7. If the impulse is applied about the axle of the top, to increase OC to OC_3, and make the axle rise to a cusp in fig. 3,

$$OC_3\,.\,K_3M_3=2OM\,.\,MK,\quad\text{with}\quad MK=K_3M_3,$$

so that $OC_3=2OM$.

Thus the axle will rise from the horizontal in fig. 1 to a cusp by the application of an axial impulse $CC_3=OC$.

8. The impulse might be applied about a vertical axis to the steady motion, making K rise vertically to K_3; and then in a cusped motion, with OC changed to OC_3, and rising to OC_2 at K_2 on the level of G_3K_3 in fig. 4,

$$G_3K_3^2-G_3K_2^2=OK_3^2-OK_2^2=C_3K_3^2=MK_3\,.\,K_3D$$
$$=2OM\,.\,MK\,(\cos\theta_2-\cos\theta_3)=2OM\,.\,MK\,\frac{K_3D}{OC_3},$$
$$MK_3\,.\,OC_3=2OM\,.\,MK=OM\,.\,MK',$$
$$\frac{MK_3}{MK'}=\frac{MC_3}{MC'}=\frac{OM}{OC_3},$$

dropping the perpendiculars K_3C_3, K′C′ on OC; and drawing the circle on the diameter OC′, with ordinate MQ′,

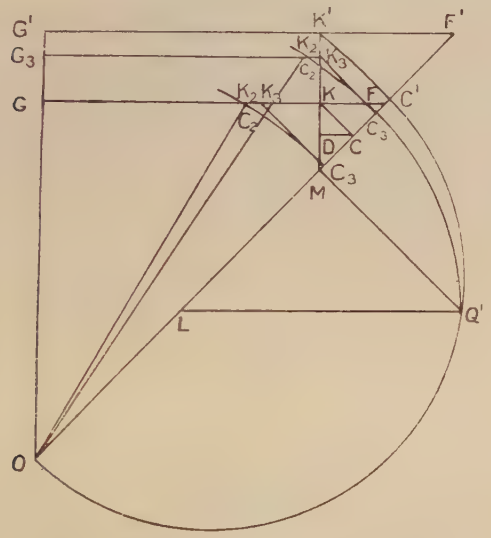
$$C_3M\,.\,C_3O=OM\,.\,MC'=MQ'^2.$$

Then if L is the midpoint of OM,

$$C_3M \cdot C_3O = LC_3^2 - LM^2 = LQ'^2, \qquad LC_3 = LQ',$$

to determine C_3.

Fig. 4.



$MK' = 2MK$. L mid-point of OM.
$MC' = 2MC$. $LC_3 = LQ'$.

9. Applied about a horizontal axis in the vertical plane GOC, the impulse will make K move horizontally to K_3, and the axle rise to a cusp at K_2, C_2, in fig. 4, if

$$C_3K_2^2 = M_3K_3 \cdot K_3D = 2OM \cdot MK \frac{K_3D}{OC},$$

$$M_3K_3 \cdot OC_3 = 2OM \cdot MK, \qquad \frac{M_3K_3}{MK} = \frac{FC_3}{FC} = 2\frac{OM}{OC_3},$$

$$FC_3 \cdot C_3O = 2OM \cdot FC - 2GK \cdot KF = G'K' \cdot K'F',$$

thence a geometrical construction may be devised for the determination of C_3 and K_3.

10. When the impulse is applied in the vertical plane GOC, the impulse vector starts out of the plane, from OK to OH, and K moves to H perpendicular to the plane GOC.

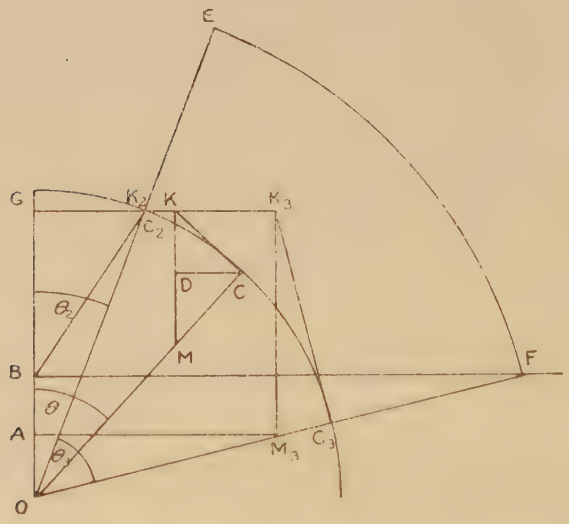
The axle then oscillates between θ_2 and θ_3,

$$\theta_2<\theta<\theta_3, \qquad \mathrm{KH}=\mathrm{A}\frac{d\theta}{dt},$$

$$\mathrm{GH}^2=\mathrm{KH}^2+\mathrm{GK}^2=\mathrm{KH}^2+\left(\frac{\mathrm{OC}-\mathrm{OG}\cos\theta}{\sin\theta}\right)^2$$
$$=2\mathrm{OM}\,.\,\mathrm{MK}\,(\mathrm{E}-\cos\theta),$$

$$\mathrm{CH}^2=\mathrm{KH}^2+\mathrm{CK}^2=\mathrm{KH}^2+\left(\frac{\mathrm{OG}-\mathrm{OC}\cos\theta}{\sin\theta}\right)^2$$
$$=2\mathrm{OM}\,.\,\mathrm{MK}\,(\mathrm{D}-\cos\theta).$$

Fig. 5.



When the axle rises to a cusp, C rises to C_2, K_2 at the level of G, where in fig. 5

$$\mathrm{KH}=0, \qquad \mathrm{OG}-\mathrm{OC}\cos\theta_2=0, \qquad \mathrm{D}=\cos\theta_2,$$

$$\mathrm{CH}^2=2\mathrm{OM}\,.\,\mathrm{MK}\,(\cos\theta_2-\cos\theta)$$
$$=2\mathrm{OM}\,.\,\mathrm{MK}\,\frac{\mathrm{KD}}{\mathrm{OC}}=2\,\frac{\mathrm{OM}}{\mathrm{OC}}\,.\,\mathrm{KC}^2$$

$$\mathrm{KH}^2=\mathrm{CH}^2-\mathrm{CK}^2=\left(2\,\frac{\mathrm{OM}}{\mathrm{OC}}-1\right)\mathrm{CK}^2,$$

giving KH the impulse.

This can be applied by hammering the rim of the bicycle wheel with a stick in a vertical blow at its highest point.

Then

$$OK_3^2 - OK_2^2 = OK_3^2 - OC_3^2 = C_3K_3^2$$
$$= 2\,OM \,.\, MK\,(\cos\theta_2 - \cos\theta_3),$$

$$M_3K_3 \,.\, K_3D_3 = 2\,OM \,.\, MK\,\frac{K_3D_3}{OC}, \quad M_3K_3 \,.\, OC = 2\,OM \,.\, MK,$$

and this determines the level of AM_3 in fig. 5, and provides a geometrical construction for the position of C_3.

For

$$OM_3 \,.\, M_3C_3 = OA \sec\theta_3 \,.\, M_3K_3 \cos\theta_3 = OA \,.\, AG,$$

so that, if B is taken in OG where AB = OA,

$$4\,OA \,.\, AG = OG^2 - BG^2 = OK_2^2 - K_2B^2\,;$$

and the circle centre K_2 and radius K_2B will cut off a length OE on $OK_2 = OF = 2\,OM_3$, and so determine the direction of OM_3C_3, the axle in the lowest position.

When OC is horizontal, as in fig. 1,

$$\cos\theta_3 = 0, \quad CH^2 = 2\,OC \,.\, CK\cos\theta_2 = 2\,CK^2, \quad KH = KC,$$

as at first; and $OC = OM$, $M_3K_3 = 2\,MK$, $\theta_2 + \theta_3 = \pi$, the axle oscillating to an equal angle above and below the horizontal.

11. When the cusp motion is pseudo-regular and in small loops, it can be projected on the tangent cone of a sphere in a series of small hypocycloidal branches, and the motion is realised as discussed in the *Principia*, Book I, section X, when the tangent cone is developed into a plane and gravity radiates from a centre.

With the axle horizontal, a necklace of brachistochrone cycloids is formed round the equator, with mean regular precession μ, fluctuating in azimuth between 0 and 2μ, with azimuth interval $2\Psi = 2\,K\kappa$, tending as the rotation R and axial impulse CR is increased to $\pi\cos\theta_2$, in a zone above the equator of angular width $\frac{1}{2}\pi - \theta_2$, and area $2\pi a^2\cos\theta_2$ on a sphere of radius a, and the number of cycloids in the necklace would be about $2\sec\theta_2$.

12. Even in the steadiest smoothest Regular Precession a close scrutiny will reveal under the slightest disturbance an almost invisible deviation from a perfect circular motion, in the shape of a progressive motion of an apse line, realised

easily with the thread and plummet; utilised by Newton to illustrate the Evection of the Moon.

Returning to the general unsteady motion of a top in § 3, and its vector impulse OH, the velocity of H imparted by the gravity couple $An^2 \sin\theta$ is horizontal and along KH perpendicular to the plane GOC. This velocity is the rate of growth of KH, added to the velocity of K carried round by the plane GOC with angular velocity $\frac{d\psi}{dt}$: so that, with $KC = A\sin\theta\frac{d\psi}{dt}$, and putting $\frac{d\theta}{dt} = Q$,

$$An^2\sin\theta = A\frac{d^2\theta}{dt^2} + GK\frac{d\psi}{dt} = A\frac{dQ}{dt} + \frac{GK\,.\,KC}{A\sin\theta},$$

with $GK = \frac{OC - OG\cos\theta}{\sin\theta}$, $KC = \frac{OG - OC\cos\theta}{\sin\theta}$,

and $\frac{dGK}{d\theta} = \frac{KC}{\sin\theta}$, $\frac{dKC}{d\theta} = \frac{GK}{\sin\theta}$,

obvious geometrically on fig. 3.

Differentiating with respect to θ, with $d\theta = Q\,dt$,

$$\frac{1}{Q}\frac{d^2Q}{dt^2} = n^2\cos\theta - \frac{GK^2 + KC^2 - GK\,.\,KC\cos\theta}{A^2\sin^2\theta}$$

$$= n^2\cos\theta - \frac{GC^2 - 3GK\,.\,KC\cos\theta}{A^2\sin^2\theta}$$

$$= \left(n^2 + \frac{3OM\,.\,ON}{A^2}\right)\cos\theta - \frac{OK^2}{A^2},$$

an exact equation.

In a state of perfect Steady Motion of Regular Precession, Q and $\frac{dQ}{dt}$ are zero, and, in dynamical units, from § 5,

$$OM\,.\,ON = A^2n^2, \qquad \frac{OM}{An} = \frac{n}{\mu}, \qquad \frac{OM}{ON} = \frac{n^2}{\mu^2} = \frac{\lambda}{l},$$

where λ is the height of the equivalent conical pendulum,

$$\lambda = g/\mu^2 = OA, \qquad l = g/n^2 = OP;$$

$$\frac{OG}{An} = \frac{n}{\mu}\cos\theta + \frac{\mu}{n}, \qquad \frac{OC}{An} = \frac{n}{\mu} + \frac{\mu}{n}\cos\theta,$$

$$\frac{GC^2}{An^2} = \frac{OK^2\sin^2\theta}{A^2n^2} = \left(\frac{n^2}{\mu^2} + 2\cos\theta + \frac{\mu^2}{n^2}\right)\sin^2\theta.$$

Then in this Steady Motion,

$$\frac{1}{Q}\frac{d^2Q}{dt^2}+m^2=0, \quad \text{where}$$

$$\frac{m^2}{n^2}=\frac{OK^2}{A^2n^2}-4\cos\theta=\frac{\lambda}{l}-2\cos\theta+\frac{l}{\lambda}=\frac{AP^2}{OA\,.\,OP},$$

$$\frac{g}{m^2}=\frac{OP^2}{AP^2}\ OA.$$

This result is exact and reached without any approximation: and the slightest disturbance will give a nutation $Q=Q_0\cos(mt+\epsilon)$, beating $m/2\pi$ times a second, and the apsidal angle, from node to node, is

$$\Psi=\mu\frac{\pi}{m}=\pi\frac{OP}{AP}.$$

In Darboux's representation of top-motion by a deformable articulated hyperboloid of the generating lines, the model is flattened into a rigid framework for Steady Motion; and KM, KN produced to double length at S, S′ will make the focal line SS′ parallel to MN; this will be revolved about the vertical line ON with constant angular velocity μ. The small nutation will be due to a slight play or backlash in the frame.

13. The same argument can be applied to the invisible oscillation of a Simple or Spherical Pendulum, or to the apsidal angle of a particle describing a horizontal circle on a smooth surface of revolution about a vertical axis.

Taken as the axis Oy, the general equations of motion of the particle are

$$\frac{1}{2}\frac{dx^2}{dt^2}+\frac{1}{2}\frac{dy^2}{dt^2}+\frac{1}{2}x^2\frac{d\psi^2}{dt^2}+gy=H \text{ (energy)},$$

$$\text{and} \quad x^2\frac{d\psi}{dt}=K \text{ (impulse)};$$

so that, with $\frac{dx}{dt}=Q$, and eliminating $\frac{d\psi}{dt}$,

$$\frac{1}{2}Q^2\left(1+\frac{dy^2}{dx^2}\right)+\frac{1}{2}\frac{K^2}{x^2}+gy=H;$$

and differentiating with respect to x, $dx=Q\,dt$, $Q\frac{dQ}{dx}=\frac{dQ}{dt}$,

$$\frac{dQ}{dt}\left(1+\frac{dy^2}{dx^2}\right)+Q\frac{dy}{dx}\frac{d^2y}{dx^2}-\frac{K^2}{x^3}+g\frac{dy}{dx}=0,$$

$$\frac{1}{Q}\frac{d^2Q}{dt^2}\left(1+\frac{dy^2}{dx^2}\right)+2\frac{dQ}{dt}\frac{dy}{dx}\frac{d^2y}{dx^2}+Q\left(\frac{d^2y}{dx^2}\right)^2$$

$$+Q\frac{dy}{dx}\frac{d^2y}{dx^2}+\frac{3K^2}{x^4}+g\frac{d^2y}{dx^2}=0,$$

exact equations.

In a state of Steady Motion in a horizontal circle,

$$Q=0, \quad \frac{dQ}{dt}=0, \quad K=\mu x^2,$$

$$g=\mu^2 x\frac{dx}{dy}=\mu^2 . NG, \quad v^2=\mu^2x^2=gx\frac{dy}{dx}=g . NV,$$

if the normal and tangent at P meet the axis Oy in G and V; and then with $\frac{dy}{dx}=\tan\theta$,

$$\frac{1}{Q}\frac{d^2Q}{dt^2}+m^2=0, \quad \frac{m^2}{\mu^2}=\left(3+NG\frac{d^2y}{dx^2}\right)\cos^2\theta;$$

θ oscillates between close limits $\alpha\pm\beta$,

$$\theta=(\alpha+\beta)\sin^2\tfrac{1}{2}mt+(\alpha-\beta)\cos^2\tfrac{1}{2}mt, \quad Q=m\beta\sin mt;$$

and the particle beats $m/2\pi$ nutations per second, synchronizing with the beat of a simple pendulum of length $\lambda=g/m^2$, where

$$\frac{1}{\lambda}=\left(\frac{3}{NG}+\frac{d^2y}{dx^2}\right)\cos^2\theta=\frac{3}{GV}+\frac{1}{PR},$$

where PR is the semi-vertical chord of curvature upward of the profile curve of the surface.

Then on a cone $\lambda=\frac{1}{3}GV$; and on a sphere or spherical pendulum, PR=NG, and to radius a,

$$\frac{a}{\lambda}=3\cos\theta+\sec\theta.$$

For a profile given by $y=cx^n$,

$$\frac{dy}{dx}=n\frac{y}{x}, \quad NG=x\frac{dx}{dy}=\frac{x^2}{ny}, \quad v^2=g . NV=gny,$$

$$\frac{d^2y}{dx^2}=\frac{n(n-1)y}{x^2}, \quad 3+NG\frac{d^2y}{dx^2}=n+2, \quad \lambda=\frac{GV}{n+2}.$$

Thus on the surface of a free vortex, where the angular impulse xv is constant of all annular elements of liquid of the same volume, x^2y is constant, and $n=-2$, $\lambda=\infty$.

On a motor or bicycle track of this shape, the steering will be easy, and a change of place can be made without difficulty or danger, as with the annular elements of liquids in the vortex volume.

A start is made with moderate velocity from the circumference of the track where the slope is slight, and the car

is steered with increasing velocity down towards the middle, where the cars can pass and repass without difficulty.

To avoid a deep hole in the sink in the middle, the profile can change to the parabola of a forced vortex, where

$$v=\mu x, \quad n=2, \quad y=\frac{x^2}{2p}, \quad \mathrm{NG}=p, \quad \lambda=\tfrac{1}{4}\mathrm{GV}=\tfrac{1}{2}\mathrm{SP}.$$

On a horizontal circle of this track of one lap to the mile, NP = 840 feet; described in two minutes at 30 miles an hour, NV = 60 feet, and $\cot\theta=\frac{\mathrm{NP}}{\mathrm{NV}}=14$, a slope of 4°.

Raise the speed to 60 miles an hour on this track, NP = 420, NV = 270, feet, and the slope is nearly 30°, the circuit of two laps to the mile made in 30 seconds.

At a speed limit of 90 miles an hour, NP = 280, NV = 540, feet; round a circle of three laps to the mile, on a slope of over 62°. The surface could then change to a paraboloid, with a flat area in the middle, where a car could come to rest.

XIII. *The Binding of Electrons by Atoms.* *By* J. W. Nicholson, *F.R.S., Fellow of Balliol College, Oxford* *.

ACCORDING to the quantum theory of atomic structure and of the emission of line spectra, the paths of the electron in the atom vary according to the particular coordinates used in the process of quantizing the separate momenta. Thus in the simple case of a hydrogen atom, containing a nucleus and one electron, we may use either spherical polar or parabolic coordinates, and the admissible orbits are entirely different in the two cases. Yet the final values of the atomic energy are the same, and consequently each method yields the same theoretical spectrum. It has been suggested that there is in fact, in every case, only one type of coordinates which can be used, when all the modifying circumstances, such as the variation of the mass of the electron with speed, are taken into account. The only problems yet solved are those in which the separation of variables, after the manner of Jacobi, can be effected, and the contention is in fact that there is, in every case, only one set of coordinates which allows this separation, when non-degenerate cases of the motion are discussed.

But it is generally believed that the atomic energy is in all cases determinate and definite. We shall show, in the

* Communicated by the Author.

first place, that this conclusion requires modification when the path extends to infinity. The hyperbolic orbits of Epstein, which have been used extensively in the interpretation of certain groups of γ rays associated with many of the chemical atoms, constitute an instance, and we shall show that they rest on a mathematical error, and that in fact it is not possible to preserve finite phase-integrals in the process of quantizing the momenta. In fact, it appears that the whole process is only applicable to finite paths, and gives no clue to the phenomena taking place during the binding of an electron which comes from a considerable distance.

In another form, the question we propose is as to whether a hyperbolic path is possible in the same way as an elliptic one. Such would, of course, be characterized by a *positive* energy W. Certain available evidence of a simple kind, apparently not hitherto noticed, is in existence. For the existence of such paths involves the existence of parabolic paths, with $W=0$. In passage from a stationary state of energy W_1 (negative) to a parabolic path taking the electron outside the atom altogether, a quantity of energy W_1 should be involved. Spectral lines given by

$$h\nu = W_n,$$

where W_n corresponds to any one of the stationary states, should thus exist. In other words, the 'limits' of spectral series should themselves be spectral lines. But there are two reasons why evidence on these lines cannot be decisive, especially when it is negative evidence. For in the first place, the values of W_n determining the limits of series are of such magnitude that only for two or three, in any case, can the corresponding lines come into the visible spectrum, and with only hydrogen atoms and charged helium atoms to test, and enormous band spectra for both elements, the test cannot readily be applied. Moreover, the probability of an electron entering the atom in a parabolic rather than a hyperbolic path is so small that any resulting lines could hardly be expected to be of visible intensity under ordinary conditions. We consider, therefore, that the question whether limits of series are themselves spectral lines, on the principles of the quantum theory, cannot, at least at this juncture, be examined in the light of experiment, and that it must remain a matter of deduction from other phenomena.

We find it necessary, as stated, to disagree with the hypothesis, explicitly indicated several times by Sommerfeld and others, and implicitly assumed at least by the remaining

writers on the quantum theory of spectra, that the energy W is always completely determinate when all the momenta are quantized. This can be disproved not only for fictitious laws of force in an atom, but for laws which must actually occur in systems with an existence, if only a temporary one.

Consider, for example, a simple doublet and an electron in orbital motion about it. Regarding the doublet as stationary, and of moment M, its external potential is

$$V = \frac{M \cos\theta}{r^2}$$

when it is situated at the origin, with its axis along the axis of z, using spherical polar coordinates. The equation of energy for an electron moving in its presence is

$$\tfrac{1}{2}m\{\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\} + \frac{Me\cos\theta}{r^2} = -W.$$

The momenta are, in the usual notation,

$$p_1 = \frac{\partial T}{\partial \dot{r}} = m\dot{r}, \qquad p_2 = \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta},$$

$$p_3 = \frac{\partial T}{\partial \dot{\phi}} = mr^2\sin^2\theta\dot{\phi},$$

so that

$$\tfrac{1}{2}m\left\{p_1^2 + \frac{p_2^2}{r^2} + \frac{p_3^2}{r^2\sin^2\theta}\right\} + \frac{Me\cos\theta}{r^2} = -W.$$

Now ϕ is a speed coordinate as usual, so that

$$p_3 = \text{const.} = n_1 h/2\pi$$

when subjected to the quantum relation, n_1 being an integer.

For the Jacobi solution, we must also take, in separating variables,

$$p_2^2 + \frac{p_3^2}{\sin^2\theta} + 2mMe\cos\theta = \beta$$

where β is constant, and

$$\tfrac{1}{2}m\left\{p_1^2 + \frac{\beta}{r^2}\right\} = -W.$$

Thus

$$p_1 = \sqrt{-2mW - \frac{\beta}{r^2}}.$$

With a positive W, the motion is not real. Thus W must be negative and the path necessarily extends to infinity. A critical value of r is $\sqrt{\frac{\beta}{-2m\mathrm{W}}}$, and the other is infinity. The phase-integral for p_1 is

$$2\int_{\sqrt{\frac{\beta}{-2m\mathrm{W}}}}^{\infty} p_1\,dr$$

which is infinite, but nevertheless independent of W. For writing

$$r=\sqrt{\frac{\beta}{-2m\mathrm{W}}}\cdot\sigma, \qquad \text{it becomes}$$

$$2\int_1^{\infty}\sqrt{-2m\mathrm{W}}\left(1-\frac{1}{\sigma^2}\right)^{\frac{1}{2}}\sqrt{\frac{\beta}{-2m\mathrm{W}}}\,d\sigma = 2\sqrt{\beta}\int_1^{\infty}d\sigma\left(1-\frac{1}{\sigma^2}\right)^{\frac{1}{2}}.$$

A finite integral is secured,—Epstein's procedure, for instance,—by using the phase-integral not for p_1, but for $p_1-(p_1)_{r=\infty}$, which in the same way yields

$$2\sqrt{\beta}\int_1^{\infty}d\sigma\left\{\left(1-\frac{1}{\sigma^2}\right)^{\frac{1}{2}}-1\right\}=2\sqrt{\beta}\left(1-\frac{\pi}{2}\right),$$

again independent of W. Now β is quantized, or expressed definitely in terms of integers already, from the phase-integral for the momentum p_2. The phase integral for p_1 can only, in this case, lead to another expression of similar type for β, but to no expression for W. It is not at all clear that the two expressions for β, also, can both be valid simultaneously.

This possibility has hitherto apparently been overlooked by authors in this subject.

No case has, however, been noticed in which W is indeterminate for a *finite* path. One very important conclusion is that the whole investigation is valid for a negatively charged atom with a distant electron.

We proceed now to discuss the possible existence of definite paths with a positive total energy and infinite extent, for a single electron around a nucleus of charge ve, situated at the origin. This is Epstein's problem, which he treats as only two-dimensional. The energy equation is

$$\tfrac{1}{2}m\left\{p_1^2+\frac{p_2^2}{r^2}+\frac{p_3^2}{r^2\sin^2\theta}\right\}-\frac{ve^2}{r}=\mathrm{W},$$

where W is positive, and represents the total energy, and the p's are the momenta.

We have thus

$$p_3 = \text{const.} = \frac{n_1 h}{2\pi},$$

$$p_2^2 + \frac{p_3^2}{\sin^2\theta} = \beta^2,$$

being clearly positive,

$$\tfrac{1}{2}m\left\{p_1^2 + \frac{\beta^2}{r^2}\right\} - \frac{ve^2}{r} = \mathrm{W}.$$

The phase-integral for p_2 is

$$n_2 h = \int p_2\, d\theta = 2\int d\theta \,.\, \sqrt{\beta^2 - \frac{p_3^2}{\sin^2\theta}},$$

the limits being the suitable values of θ for which $p_2 = 0$. The factor 2 represents the double journey in this co-ordinate,

$$\sin\psi = \frac{p_3}{\beta},$$

where ψ is one of the limits, and the other admissible value, for a real integral, is $\pi - \psi$. Thus

$$n_2 h = 2\int_{\psi}^{\pi-\psi} d\theta\, \sqrt{\beta^2 - p_3^2}/\sin^2\theta$$

$$= 2\beta \int_1^{\frac{\beta}{p_3}} \frac{dw}{w}\sqrt{\frac{w-1}{\frac{\beta}{p_3} - w}} \qquad \left(\text{with } \sin\theta = \frac{p_3}{\beta}\sqrt{w}\right).$$

Write

$$w = \sin^2\omega + \frac{\beta^2}{p_3^2}\cos^2\omega,$$

and we have

$$n_2 h = 4\beta \cdot \frac{\beta^2 - p_3^2}{p_3^2}\int_0^{\pi/2} \frac{\cos^2\omega\, d\omega}{\sin^2\omega + \frac{\beta^2}{p_3^2}\cos^2\omega},$$

or with $\tan\omega = t$,

$$n_2 h = \frac{4\beta}{p_3^2}(\beta^2 - p_3^2)\int_0^{\infty} \frac{dt}{(1+t^2)\left(\frac{\beta^2}{p_3^2} + t^2\right)}$$

$$= 4\beta\left\{\tan^{-1} t - \frac{p_3}{\beta}\tan^{-1}\frac{p_3 t}{\beta}\right\}_0$$

$$= 2\pi\beta\left\{1 - \frac{p_3}{\beta}\right\} = 2\pi(\beta - p_3),$$

whence

$$\beta = \frac{n_2 h}{2\pi} + p_3 = (n_1 + n_2)\frac{h}{2\pi},$$

these integers being thus additive, in the usual way.

The phase-integral for p_1 is

$$n_3 h = \int dr \sqrt{2m\,\mathrm{W} + \frac{2mve^2}{r} - \frac{\beta^2}{r^2}},$$

if we seek to quantize p_1 as it stands. The limits would then be a positive value of r and infinity, for half the path, and the integral would be infinite. But it is clearly necessary to suppose that when the electron is at infinity, out of range of action of the nucleus, it should not be subject to a quantum relation, so that $(p_1)_{r=\infty}$ is not affected by the rule, and only the variable part

$$p_1 - (p_1)_\infty$$

is so affected. Yet this question of quantizing p_1 presents some difficulties in whatever way it is suggested that it should be effected, and we consider that Epstein's discussion of the matter is very incomplete and not logically justifiable in its mathematical procedure. We shall thus consider various alternatives which may give a finite phase-integral.

Now the actual r-path is not a passage from $r=\alpha$ (say) to $r=\infty$ and back, and the phase-integral is not twice the definite integral between these limits. The electron goes from a limiting radius to infinity, and back to the same radius elsewhere, and the passage through infinity distinguishes this phase-integral from those which occur in the other coordinates.

We must, of course, also remember that the sign of p_1 depends upon the part of the path concerned,—whether the electron is departing or returning. The critical value of r is the positive root of

$$2m\mathrm{W} + \frac{2mve^2}{r} - \frac{\beta^2}{r^2}$$

or

$$\frac{1}{r} = \frac{mve^2 + \sqrt{m^2v^2e^4 + 2m\mathrm{W}\beta^2}}{\beta^2} = \frac{1}{\alpha}\ (\text{say}).$$

Writing, generally, with a new variable ϕ,

$$\frac{1}{r} - \frac{mve^2}{\beta^2} = \sqrt{\frac{m^2v^2e^4 + 2m\mathrm{W}\beta^2}{\beta^4}}\cos\phi,$$

we have $\phi=0$ in the critical position (perihelion, in the usual terminology), and

$$\cos\phi=-\frac{mve^2}{\beta^2}\Big/\sqrt{\frac{m^2v^2e^4+2mW\beta^2}{\beta^4}}=\cos\eta\ (\text{say}),$$

when $r=\infty$.

What is required for the correct evaluation of the phase-integral is a *continuous* variable which shall change in one direction,—and thus give a definite integral,—as r goes to infinity and returns, the sign of p_1 being automatically taken into account,—or the sign of $p_1-(p_1)_\infty$ when $(p_1)_\infty$ is not zero as in a parabolic path. The new variable ϕ has this property, and ranges from zero to 2π as r goes through its changes. We have denoted its value, when $r=\infty$, by η above, where η is evidently an obtuse angle.

The phase-integral for p_1 alone would be

$$n_3h=\int_{\phi=0}^{2\pi}dr\,\sqrt{2mW+\frac{2mve^2}{r}-\frac{\beta^2}{r^2}}$$

(the square root being properly interpreted in different regions) where

$$\frac{1}{r}=\frac{mve^2}{\beta^2}+\frac{1}{\beta^2}\sqrt{m^2v^2e^4+2mW\beta^2}\cos\phi$$

$$0=\frac{mve^2}{\beta^2}+\frac{1}{\beta^2}\sqrt{m^2v^2e^4+2mW\beta^2}\cos\eta,$$

and we find

$$dr=\frac{\sin\phi\,d\phi}{(\cos\phi-\cos\eta)^2}\,\frac{\beta^2}{q},\qquad q=\sqrt{2mW+m^2v^2e^4},$$

$$\sqrt{2mW+\frac{2mve^2}{r}-\frac{\beta^2}{r^2}}=\frac{q}{\beta}\sin\phi.$$

If the integration were continuous throughout,—as assumed by Epstein,—we should thus have

$$n_3h=\beta\int_0^{2\pi}\frac{\sin^2\phi\,d\phi}{(\cos\phi-\cos\eta)^2}$$

$$=2\beta\int_0^{\pi}\frac{\sin^2\phi\,d\phi}{(\cos\phi-\cos\eta)^2},$$

which is an infinite integral, as would be expected.

If we merely quantized over the finite part of the hyperbola,—another possible suggestion,—we should have

$$n_3 h = \beta \left\{ \int_0^{\eta} + \int_{2\pi-\eta}^{2\pi} \right\} \frac{\sin^2\phi}{(\cos\phi - \cos\eta)^2} d\phi$$

$$= 2\beta \int_0^{\eta} \frac{\sin^2\phi \, d\phi}{(\cos\phi - \cos\eta)^2},$$

which is again infinite.

The nature of the first infinity merits a remark, however, for it is independent of η and therefore of W. For

$$I = \int_0^{\pi} \frac{\sin^2\phi \, d\phi}{(\cos\phi - \cos\eta)^2} = \left[\frac{\sin\phi}{\cos\phi - \cos\eta}\right] - \int_0^{\pi} \frac{\cos\phi \, d\phi}{\cos\phi - \cos\eta}$$

$$= \left[\frac{\sin\phi}{\cos\phi - \cos\eta}\right] - \pi - \cos\eta \int_0^{\pi} \frac{d\phi}{\cos\phi - \cos\eta}.$$

The principal value of the last integral is well known to be zero, for all values of η, so that the last term is zero. Our equation would be

$$n_3 h = -2\pi\beta + \left[\frac{\sin\phi}{\cos\phi - \cos\eta}\right],$$

where the principal value of the bracket must be taken, *i. e.* it is to be interpreted as

$$\mathrm{Lt}_{\epsilon\to 0} \left\{ \left[\frac{\sin\phi}{\cos\phi - \cos\eta}\right]_0^{\eta-\epsilon} + \left[\frac{\sin\phi}{\cos\phi - \cos\eta}\right]_{\eta+\epsilon}^{\pi} \right\}.$$

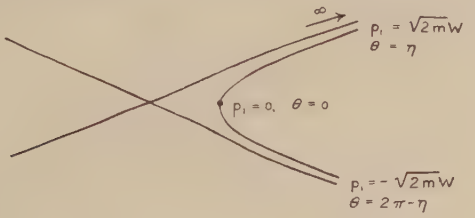
This becomes

$$\mathrm{Lt}_{\epsilon\to 0} \left\{ \frac{\sin\eta}{\epsilon \sin\eta} + \frac{\sin\eta}{\epsilon \sin\eta} \right\} = \mathrm{Lt} \cdot \frac{1}{2\epsilon},$$

which, though infinite, is an infinity independent of η and therefore of W. We have another aspect of the indeterminateness of W for such paths.

Our fundamental objection to Epstein's mode of integration may now be introduced. He integrates $p_1 - (p_1)_\infty$, and not p_1, but this fact does not affect the question. For as ϕ ranges between 0 and 2π, if $p_1 = f(\phi)$, we have p_1 varying continuously with θ, and remaining positive, till $\theta = \eta$. Then p_1 becomes $-f(\phi_1)$ when $\phi = 2\pi - \phi_1$ on the return journey after $\phi = 2\pi - \eta$. Between $\phi = \eta$ and $\phi = 2\pi - \eta$, the value of r should be infinite, and p_1 changes from $\sqrt{2m\mathrm{W}}$ to $-\sqrt{2m\mathrm{W}}$, as in the figure.

The variation of p_1 between $\pm\sqrt{2mW}$ at infinity is the source of trouble, and it takes place while $\frac{1}{r}=0$.



Epstein takes twice the integral from $\phi=0$ to $\phi=\pi$, but according to the substitution formula, r is *negative* when ϕ goes from η to π, and negative values of r are clearly not permissible. A suitable integration for the infinite region cannot in fact be effected, and any supposition of a suitable variable in place of ϕ, for the change of p_1 at ∞ from $\sqrt{2mW}$ to $-\sqrt{2mW}$, would be entirely arbitrary,—but as it could not lead to a finite phase-integral, we pursue the matter no further.

These considerations, nevertheless, have considerable force when, thrown back as we now are upon the necessity, if the quantum theory is applicable, of using $p_1-(p_1)_\infty$, we attempt to quantize this.

We have, when $\theta=\eta$

$$(p_1)_\infty = \sqrt{2mW} = \frac{q}{\beta}\sin\eta,$$

where $q=\sqrt{m^2v^2e^4+2mW\beta^2}$ as before.

And when $\phi=2\pi-\eta$,

$$(p_1)_\infty = \frac{q_1}{\beta}\sin(2\pi-\eta) = -\frac{q}{\beta}\sin\eta.$$

From $\phi=0$ to $\phi=\eta$,

$$p_1-(p_1)_\infty = \frac{q}{\beta}(\sin\phi-\sin\eta).$$

From $\phi=2\pi-\eta$ to $\phi=2\pi$,

$$p_1-(p_1)_\infty = +\frac{q}{\beta}(\sin\phi+\sin\eta),$$

and from $\phi=\eta$ to $\phi=2\pi-\eta$,

$$p_1-(p_1)_\infty=0.$$

With the value of dr, the phase-integral

$$n_3 h = \int_{\theta=0}^{2\pi} dr(p_1 - (p_1)_\infty)$$

breaks into three parts, thus

$$n_3 h = \beta \int_0^{\eta} \frac{\sin\phi\,(\sin\phi - \sin\eta)d\phi}{(\cos\phi - \cos\eta)^2} + \beta \int_{\eta}^{2\pi-\eta} 0 \,.\, d\phi$$
$$+ \beta \int_{2\pi-\eta}^{2\pi} + \frac{\sin\phi\,(\sin\phi + \sin\eta)\,d\phi}{(\cos\phi - \cos\eta)^2}$$
$$= 2\beta \int_0^{\eta} \frac{\sin\phi(\sin\phi - \sin\eta)\,d\phi}{(\cos\phi - \cos\eta)^2} + 2\beta \int_{\eta}^{\pi} 0 \,.\, d\phi$$

by a simple transformation.

Finally, the only accurate phase-integral is

$$n_3 h = 2\beta \int_0^{\eta} \frac{\sin\phi\,(\sin\phi - \sin\eta)}{(\cos\phi - \cos\eta)^2}\, d\phi,$$

while Epstein gives, in our notation,

$$n_3 h = 2\beta \int_0^{\pi} \frac{\sin\phi\,(\sin\phi - \sin\eta)}{(\cos\phi - \cos\eta)^2}\, d\phi,$$

the part of his range from η to π involving a meaningless negative value of r, and violating $p_1 = (p_1)_\infty$ though the moving electron is at infinity. The principal value of Epstein's integral is, using the indefinite integral for the function in the form, readily obtained by parts,

$$\int \frac{\sin\phi\,(\sin\phi - \sin\eta)}{(\cos\phi - \cos\eta)^2}\, d\phi$$
$$= \frac{\sin\phi - \sin\eta}{\cos\phi - \cos\eta} - \phi + \cot\eta \,.\, \log_e \left\{ \frac{\sin\frac{\eta - \phi}{2}}{\sin\frac{\eta + \phi}{2}} \right\}$$

of the type

$$n_3 h = 2\beta\pi \left\{ -1 + \frac{2}{\pi \sin\eta} \right\}$$

or

$$\frac{2}{\pi \sin\eta} = \frac{(n_1 + n_2 + n_3)}{n_1 + n_2},$$

and ultimately

$$W = \frac{2\pi^2 m e^4 \nu^2}{h^2} \,.\, \frac{1}{\frac{\pi}{4}\,(n_1 + n_2 + n_3)^2 - (n_1 + n_2)^2},$$

—generalized from his value which relates only to a *plane*

hyperbola. We have the sum $n_1 + n_2$ of the angular quanta in place of his single integer.

But this formula, with all the applications he makes to characteristic γ radiation, is not tenable, as resting on a mathematical error. Its apparent success appeared at one time to the writer to justify it as an empirical formula, in spite of his independent investigation, outlined above, indicating the impossibility of quantizing such orbits. Close examination, however, of the calculations of γ radiation and so forth made it clear that they were in several cases illusory, and determined more by order of magnitude than by the nature of the formula.

There is one convincing argument against the formula, however. It should give an emission spectrum for all values of n_1, n_2, n_3 and m_1, m_2, m_3 making

$$-\mathrm{W}(m_1, m_2, m_3) + \mathrm{W}(n_1, n_2, n_3)$$

positive. This can be tested in great numerical detail on the spectrum of a hydrogen atom, and the test fails entirely. No spectrum line is found,—in the secondary hydrogen spectrum,—in any of the assigned positions. Thus the formula really fails as an empirical one.

We have seen above that it must be replaced by

$$n_3 h = 2\beta \int_0^{\eta} \frac{\sin\phi\,(\sin\phi - \sin\eta)}{(\cos\phi - \cos\eta)^2}\, d\phi$$

$$= 2\beta \left[\frac{\sin\phi - \sin\eta}{\cos\phi - \cos\eta} - \phi + \cot\eta \log_e \left\{ \frac{\sin\frac{\eta - \phi}{2}}{\sin\frac{\eta + \phi}{2}} \right\} \right]_0^{\eta}$$

which is logarithmically infinite.

The attempt to obtain a finite phase-integral, in this manner, in fact fails, and we must give up the hypothesis that even the variable part of p_1 can be quantized for the infinite path.

It is not difficult to see that this conclusion is general for any infinite path which is possible for an electron about a physically existent atom, whose nucleus can always be regarded, for the present purpose, as a superposition of free charges and a set of doublets. We have demonstrated the result for a single free charge, and previously for sets of doublets. Further analysis of the more general case does not seem necessary, and could readily be supplied by the reader.

Our conclusion must be as follows :—

A determinate and finite value of W cannot be obtained for an electron moving about any atomic nucleus, if the path involved takes the electron to infinity.

XIV. *Theoretical Aspects of the Neon Spectrum.* By Laurence St. C. Broughall *.

THE object of this paper is to attempt to explain the spectrum of neon in a manner somewhat similar to that used by Bohr † in his explanation of the reason for the existence of the Balmer series in the hydrogen spectrum.

The principle on which this hypothesis rests is that when an electron rotates in a fixed orbit it does not radiate energy, although the principles of electrodynamics state that it should; if, however, the electron changes from one orbit to another, then energy is emitted, provided that the kinetic energy of the electron is less in the second orbit than in the first.

In order to account for the spectrum, it is assumed that the energy emitted is numerically equal to the product of the frequency of the spectral line produced and the quantum constant. We thus obtain the equation

$$E = nh,$$

where E = energy emitted, n = frequency of the resulting radiation, and h = quantum constant.

In the case of hydrogen, it was assumed that the orbit of the electron was circular, and then the attractive force between nucleus and the electron due to their equal and opposite charges was balanced by the centrifugal force of the electron due to its rotation about an axis passing through the nucleus.

The energy of the electron can thus be found for any radius of orbit. When the electron changes its orbit, it moves to one with a radius which is an exact multiple of the radius of the original orbit. In this manner the change of energy due to a change of orbit can be found, and then, using the equation given above, it was shown by Bohr how the constant of the Balmer series could be found; and the value so obtained agreed extremely well with that found by experiment.

In the case of neon, we are dealing with an atom which contains more than one electron; and since the atomic number is 10, it follows that if the atom is to be neutral, then there must be 10 electrons present to annul the excess of 10 positive charges in the nucleus.

* Communicated by the Author.

† Phil. Mag. vol. xxix. p. 332.

In order, therefore, to study the atom, it is essential that the electrons be given definite position relative to one another. This has been undertaken by Langmuir *, and there is considerable evidence in favour of the postulate that eight of the electrons arrange themselves at the corners of a cube at the centre of which the nucleus is situated. The other two electrons are imagined to lie within this cube, probably on a line joining the mid-points of any pair of opposite sides. If we make use of this hypothesis, and further are in possession of data which will allow us to find the length of the diagonal of the electron cube, then it was shown by the author † that it is possible to calculate the angular velocities of the electrons about the nucleus.

Since the determination of the spectral lines is an extension of the matter given in that paper, it will be advisable here to state the principles on which the calculations of the electron frequencies depend.

It has already been stated that the two inner electrons will probably lie on a line joining the mid-points of any pair of opposite sides. If this be the case, then the electrical forces acting on the outer electrons due to the other electrons in the outer shell, and to the two inner electrons, will be the same whichever electron we take, provided that the two inner electrons are equidistant from and on opposite sides of the nucleus.

The next consideration was the axes of revolution. As before, it was desired if possible to get the forces acting on the outer electrons due to centrifugal action the same for all of the electrons. If we take as axes the three lines which pass through the mid-points of the three opposite pairs of sides of the electron cube respectively, then the above condition will be satisfied. In the diagram the axes of revolution are illustrated by XX′, YY′, and ZZ′. The inner electrons being on the axis XX′ will only rotate about two axes. It is, of course, quite immaterial which axis the inner electrons lie upon. The forces acting on any outer electron were then considered, and were taken along the three sides of the cube which meet at the point where the electron is situated. Now, since the electron must be in equilibrium, so the force along each of these lines due to electrical attraction and repulsion and also due to the motion in a circular orbit must be equal to zero.

* General Electric Review, 1919.
† Phil. Mag. Feb. 1922.

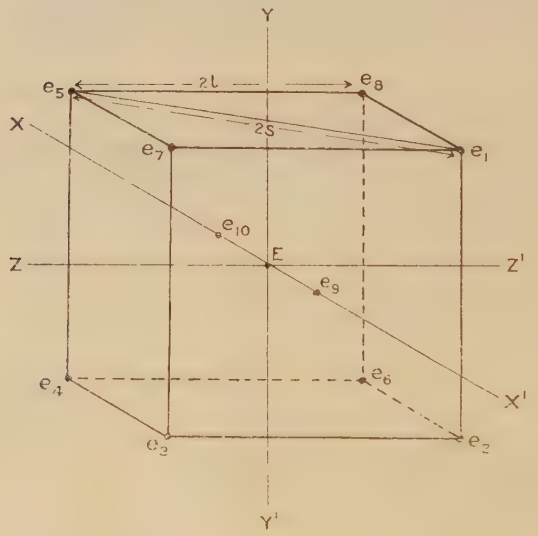
In this manner three equations were obtained, namely :

$$\frac{39e^2l}{4c^3} = \frac{e^2}{4l^2} + \frac{e^2l}{2s^3} + \frac{e^2(r+l)}{[(r+l)^2+s^2]^{3/2}} - \frac{e^2(r-l)}{[(r-l)^2+s^2]^{3/2}} + m\omega_1{}^2l + m\omega_2{}^2l, \quad . \quad . \quad . \quad \text{(I.)}$$

$$\frac{39e^2l}{4c^3} = \frac{e^2}{4l^2} + \frac{e^2l}{2s^3} + \frac{e^2l}{[(r+l)^2+s^2]^{3/2}} + \frac{e^2l}{[(r-l)^2+s^2]^{3/2}} + m\omega_1{}^2l + m\omega_3{}^2l, \quad . \quad . \quad . \quad \text{(II.)}$$

$$\frac{39e^2l}{4c^3} = \frac{e^2}{4l^2} + \frac{e^2l}{2s^3} + \frac{e^2l}{[(r+l)^2+s^2]^{3/2}} + \frac{e^2l}{[(r-l)^2+s^2]} + m\omega_2{}^2l + m\omega_3{}^2l. \quad . \quad . \quad . \quad \text{(III.)}$$

Fig. 1.



Another equation can be obtained by considering the forces acting on either of the inner electrons along the line joining the two inner electrons. Equating the forces to zero, we found that

$$\frac{10e^2}{r^2} = \frac{4e^2(r+l)}{[(r+l)^2+s^2]^{3/2}} + \frac{4e^2(r-l)}{[(r-l)^2+s^2]^{3/2}} + \frac{e^2}{4r^2} + m\omega_1{}^2r + m\omega_2{}^2r. \quad . \quad . \quad \text{(IV.)}$$

In the above equations, e = charge on an electron ;

r = radius of orbit described by the two inner electrons; ω_1 = angular velocity about YY′; ω_2 = angular velocity about ZZ′; ω_3 = angular velocity about XX′. $l = \frac{1}{2}$ length side of electron cube; $s = \frac{1}{2}$ surface diagonal of the electron cube; $c = \frac{1}{2}$ diagonal of cube. 's' and 'l' are of course functions of 'c,' and if the latter is known, then 's' and 'l' can be found. m = mass of an electron when its velocity is small compared with that of light.

From these equations it was shown by the author that an equation involving only 'r,' 'c,' 'l,' and 's' could be found. In the paper mentioned, 'l' and 's' were not expressed as functions of 'c,' but expressing them as such, since $l = \frac{\sqrt{3}c}{3}$ and $s = \frac{\sqrt{6}c}{3}$, we obtain the equation

$$\frac{4{\cdot}34}{c^2} = \frac{r+{\cdot}577c}{\left[(r+{\cdot}577c)^2+{\cdot}667c^2\right]^{3/2}}\left\{1-\frac{2{\cdot}308c}{r}\right\} - \frac{r-{\cdot}577c}{\left[(r-{\cdot}577c)^2+{\cdot}667c^2\right]^{3/2}}\left\{1+\frac{2{\cdot}308c}{r}\right\}+\frac{5{\cdot}63c}{r^3}. \quad \ldots \quad \text{(V.)}$$

In a recent article it was shown by Prof. W. L. Bragg* that the diameter of the neon atom could be found by an inspection of the diameters of the atoms of elements whose atomic numbers were near that of neon. It is impossible to measure the radius of the neon atom directly, since it forms no chemical compounds. The value obtained was very much smaller than that found by gas measurements, and the former is considered by Bragg to be the distance between the electrons in the atom—that is to say, is equal to '$2c$.' The value obtained by Chapman† from gas measurements is, however, the diameter presented by the molecule when in collision with other molecules. The difference is due to the fact that in molecular collisions in the gaseous state the outer electrons of the molecules do not come into contact‡. Using Bragg's value we have $2c = 1{\cdot}30 \times 10^{-9}$ cm., and on substituting in equation (V.) we have a means of obtaining the fundamental value of 'r.'

An inspection of equations (II.) and (III.) shows at once that $\omega_1 = \omega_2$ and equations may be obtained for ω_1 and ω_3.

* Phil. Mag. vol. xl. p. 169.
† Trans. Roy. Soc. A. vol. 216, p 279.
‡ Rankine, Proc. Roy. Soc. vol. xcviii. p. 360.

These equations take the form

$$\frac{e^2}{2mr}\left\{\frac{39}{4r^2}-4\left[\frac{r+l}{\{(r+l)^2+s^2\}^{3/2}}+\frac{r-l}{\{(r-l)^2+s^2\}^{3/2}}\right]\right\}=\omega_1^2, \quad \text{(VI.)}$$

$$\frac{e^2 r}{ml}\left\{\frac{1}{[(r-l)^2+s^2]^{3/2}}-\frac{1}{[(r+l)^2+s^2]^{3/2}}\right\}=\omega_1^2-\omega_3^2. \quad \text{(VII.)}$$

Using a slightly different value for 'm' from that used in the previous paper, we obtain the following values :—

r.	$\omega_1(=\omega_2)$.	ω_3.
$6{\cdot}19\times10^{-9}$ cm.	$6{\cdot}034\times10^{16}$ rad/sec.	$4{\cdot}290\times10^{16}$ rad./sec.,

'm' being equal to $9{\cdot}005\times10^{-28}$ grm. and $e=4{\cdot}774\times10^{-10}$ E.S.U.

These figures refer to the neon atom when in its normal state. There is some doubt as to whether they apply without modification in the gaseous state, but certain assumptions are made later in this paper which leads one to the conclusion that if the atom is larger under natural conditions, then the only result will be the elimination of certain spectral lines in the ultra-violet. When the atoms of the neighbouring elements were submitted to measurement, they constituted a solid body ; it is, therefore, quite conceivable that modification will occur if the element becomes gaseous.

In order to explain the nature of the spectral lines, we have to consider the change of energy due to a change of orbit, energy being emitted when the orbit increases in diameter.

Bohr, as already stated, imagined in the case of hydrogen that the radius of the orbit increased by constant multiples of the radius of the initial orbit. To adopt such a plan in the case of neon would lead to the emission of spectral lines of a frequency which would only give ultra-violet lines under reasonable circumstances. Further, there is no reason why the increment should be of such a nature, and the hypothesis used in our case is that the spherical shell formed by the inner electrons increases in radius until the shell has a radius equal to that of the initial outer shell of electrons. In order that equilibrium may remain, it is essential that the outer shell also expands to an extent which can be calculated from equation (V.).

The initial increment is of the nature of 3×10^{-10} cm. This process of expansion continues again and again, the inner electrons always occupying the orbit previously occupied by the outer electrons.

Having thus found a definite value for 'r' corresponding to a definite value for 'c,' we are now in a position to calculate the values of the angular velocities about the several axes. To do this, equation (VI.) is first applied; and having found 'ω_1^2,' the value obtained is substituted in equation (VII.), and ω_3^2 obtained. In calculating the value of 'c,' use is made of the fact that the ratio $c:r$ is practically constant; such ratio values are used in Table II. The closeness of the ratio figures and the absolute figures is shown in the appended table.

TABLE I.

r.	c (ratio method).	c (from eqn. V.).
$6{\cdot}50\times10^{-9}$ cm.	$6{\cdot}83\times10^{-9}$ cm.	$6{\cdot}83\times10^{-9}$ cm.
12·33 „	12·95 „	12·95 „
15·01 „	15·77 „	15·77 „

It will be seen that the agreement of the figures is so nearly exact as to warrant their use, remembering that our fundamental value of 'c' has not been obtained experimentally.

Table II. shows in columns II. and III. the values of 'c' and 'r,' and in columns IV. and V. the values of ω_1^2 and ω_3^2 respectively.

TABLE II.

I. No.	II. Radius Outer Orbit = c.	III. Radius Inner Orbit = r.	IV. ω_1^2.	V. ω_3^2.
1 ...	$6{\cdot}50\times10^{-9}$ cm.	$6{\cdot}19\times10^{-9}$ cm.	$36{\cdot}41\times10^{32}$	$18{\cdot}40\times10^{32}$
2 ...	6·83 „	6·50 „	31·41 „	15·84
3 ...	7·17 „	6·83 „	27·08 „	13·65 „
4 ...	7·54 „	7·17 „	23·36 „	11·78 „
5 ...	7·92 „	7·54 „	20·15 „	10·16 „
6 ...	8·32 „	7·92 „	17·38 „	8·768 „
7 ...	8·74 „	8·32 „	15·00 „	7·568 „
8 ...	9·18 „	8·74 „	12·93 „	6·528 „
9 ...	9·64 „	9·18 „	11·15 „	5·620 „
10 ...	10·13 „	9·64 „	9·625 „	4·860 „
11 ...	10·64 „	10·13 „	8·299 „	4·182 „
12 ...	11·18 „	10·64 „	7·160 „	3·613 „
13 ...	11·74 „	11·18 „	6·170 „	3·109 „
14 ..	12·33 „	11·74 „	5·326 „	2·684 „
15 ...	12·95 „	12·33 „	4·602 „	2·322 „
16 ...	13·60 „	12·95 „	3·970 „	2·002 „
17 ...	14·29 „	13·60 „	3·428 „	1·729 „
18 ...	15·01 „	14·29 „	2·954 „	1·489 „
19 ...	15·77 „	15·01 „	2·550 „	1·286 „
20 ...	16·57 „	15·77 „	2·201 „	1·113 „
21 ...	17·41 „	16·57 „	1·897 „	·9581 „

Let us now consider the energy of an electron in the outer shell when the diameter of the shell $=2c_1$.

Let the angular velocities about the axes YY′ and ZZ′ $=W_1$.

Let the angular velocity about axis XX′$=W_3$.

Since the diameter of the shell is equal to '$2c_1$,' it follows that the radius of the electron orbit $=S_1$ where $S_1=c_1\sqrt{6}/3$.

Using the above notation and remembering further that the kinetic energy of a particle describing a circular path of radius 'R' with an angular velocity 'W' is equal to $\frac{1}{2}MR^2W^2$, where M is the mass of the particle, we find that the kinetic energy due to rotation about the axis XX′

$$=\tfrac{1}{2}mS_1^2W_3^2,$$

and the kinetic energy due to rotation about the axis YY′ plus that due to rotation about ZZ′

$$=mS_1^2W_1^2 \quad \text{since} \quad W_1=W_2.$$

Therefore the total kinetic energy of the particle is equal to

$$\tfrac{1}{2}mS_1^2\{2W_1^2+W_3^2\}.$$

In the case of an electron in the inner shell where the radius of the orbit $=R_1$, we have the kinetic energy of the electron due to its rotation about the axes YY′ and ZZ′$=E_1$ where

$$E_1=mR_1^2W_1^2.$$

Now let the inner shell expand until it occupies the space previously occupied by the outer shell—that is to say, until $R_1=C_1$, then kinetic energy in new orbit $=E_2$.

Where $E_2=mC_1^2W_{11}^2$, 'W_{11}' being the new angular velocity about the axes YY′ and ZZ′, the change of energy, $=E_1-E_2$,

$$=m(R_1^2W_1^2-C_1^2W_{11}^2).$$

Meanwhile the outer electrons have moved further away from the nucleus, and now the outer shell has a radius $=C_2$, and the orbit of the electrons is now S_2.

Therefore the energy in the new position is equal to

$$\tfrac{1}{2}mS_2^2[2W_{11}^2+W_{13}^2].$$

Where 'W_{13}' is the new angular velocity about the axis

XX′, the change of energy due to change of orbit is therefore equal to

$$\tfrac{1}{2}m\{S_1^2[2W_1^2+W_3^2]-S_2^2[2W_{11}^2+W_{13}^2]\}.$$

Now, by Bohr's assumptions we have the equation

$$\text{Energy Emitted} = \text{Frequency} \times h.$$

Now, the frequency of a light-wave $=c/\lambda$ where 'c' is the velocity of light and 'λ' is the wave-length.

Therefore

$$\text{Energy Emitted} = \epsilon = \frac{ch}{\lambda},$$

or

$$\lambda = \frac{ch}{\epsilon}.$$

Now, in the case of an inner electron

$$\epsilon = m(R_1^2W_1^2 - C_1^2W_{11}^2);$$

$$\therefore \quad \lambda = \frac{ch}{m(R_1^2W_1^2 - C_1^2W_{11}^2)},$$

giving a series of lines for different values of R.

In the case of the outer electrons,

$$\lambda = \frac{2ch}{m\{S_1^2(2W_1^2+W_3^2)+S_2^2(2W_{11}^2+W_{13}^2)\}},$$

giving a second series of spectral lines.

Table III. shows the energies corresponding to definite radii. Column II. shows the energy content of an inner electron on the left, and that of an outer electron on the right. Column III. shows the change of energy, and column IV. shows the wave-lengths of the spectral lines produced. It should be stated here that the energy under consideration is the energy of one electron and not of the whole shell. It has been stated that there is a possibility of the atom not being in its normal condition to begin with, owing to its gaseous condition. If, however, it has expanded, then instead of starting with an atom whose diameter is $1{\cdot}30\times10^{-8}$ cm., we start with one whose radius is in all probability equal to one of the radii given in Table II. If this is so, then the only change produced will consist of the elimination of some of the lines of higher frequency.

TABLE III.

I.	II *a*.	II *b*.	III *a*.	III *b*.	IV *a*.	IV *b*.
No.	Energy per electron.		Energy Difference (ergs).		Wave-length of Spectral Line.	
	Inner.	Outer.	Inner.	Outer.	Inner.	Outer.
1 ...	$1{\cdot}255\times10^{-10}$	$1{\cdot}157\times10^{-10}$				
2 ...	1·195 „	1·101 „	$6{\cdot}0\times10^{-12}$	$5{\cdot}6\times10^{-12}$	3272 A°	3505 A°
3 ...	1·137 „	1·048 „	5·8 „	5·3 „	3386 „	3704 „
4 ...	1·082 „	·9970 „	5·5 „	5·1 „	3570 „	3850 „
5 ...	1·030 „	·9491 „	5·2 „	4·79 „	3776 „	4100 „
6 ...	·9801 „	·9035 „	4·96 „	4·56 „	3959 „	4305 „
7 ...	·9337 „	·8604 „	4·67 „	4·31 „	4205 „	4555 „
8 ...	·8886 „	·8187 „	4·51 „	4·17 „	4353 „	4709 „
9 ...	·8457 „	·7793 „	4·29 „	3·94 „	4576 „	4983 „
10 ...	·8055 „	·7423 „	4·02 „	3·70 „	4884 „	5307 „
11 ...	·7665 „	·7063 „	3·90 „	3·60 „	5034 „	5454 „
12 ...	·7297 „	·6723 „	3·68 „	3·40 „	5335 „	5775 „
13 ...	·6940 „	·6396 „	3·57 „	3·27 „	5499 „	6005 „
14 ...	·6609 „	·6089 „	3·31 „	3·07 „	5932 „	6396 „
15 ...	·6298 „	·5803 „	3·11 „	2·86 „	6313 „	6864 „
16 ...	·5995 „	·5525 „	3·03 „	2·78 „	6480 „	7063 „
17 ...	·5709 „	·5260 „	2·86 „	2·65 „	6864 „	7409 „
18 ...	·5430 „	·5003 „	2·79 „	2·57 „	7037 „	7640 „
19 ...	·5172 „	·4766 „	2·58 „	2·37 „	7610 „	8285 „
20 ...	·4927 „	·4540 „	2·45 „	2·26 „	8013 „	8688 „
21 ...	·4689 „	·4321 „	2·38 „	2·19 „	8249 „	8966 „

Owing to the complexity of the neon spectrum, it would be useless to attempt to compare our calculated lines with those found by experiment; indeed, it would be deleterious to attempt such a comparison, since the impression would be given that there is a definite line in the spectrum which corresponds to one of our calculated lines. Emphasis may only be laid upon the fact that our series produce lines in the visible part of the spectrum, which do not compare unfavourably with those obtained by experiment. Reference to Table III. will show at once that only forty lines have been determined between $\lambda = 3272$ Å and $\lambda = 8966$ Å,

whereas there are many more lines in existence. These lines can only be explained by the fact that when the spectrum is obtained, large numbers of ionized atoms exist, and under such circumstances our fundamental equations no longer hold.

II. *Ionization in the Neon Atom.*

The discussion of the properties of ionized atoms is very complex, since the possible degrees and modes of ionization are very numerous. The first case which comes under consideration is the atom which has lost one electron, thus leaving an excess of one positive charge.

It is very probable that one of the outer electrons will be removed, thus leaving seven electrons in the outer shell. Now, it seems probable that the angular momentums of the remaining electrons will suffer no change, the light produced during ionization being due solely to the change of energy of the electron suffering removal. We are thus left with seven electrons, each possessing the same angular velocity. It is a matter of considerable difficulty to arrange these electrons, and it is impossible to arrange them on a spherical surface without the force acting on an electron varying with the electron taken. We are therefore obliged to separate them into different shells.

Now, since the angular momentums of our seven outer electrons are the same, it follows that, if they are not on the the same spherical surface, then they must be in motion relative to one another. Under such conditions the positions of the electrons will vary with time.

Owing to the complexity of such a case, it seems impossible to treat the case mathematically without more experimental evidence. There are further atoms present which have been ionized to a greater extent, thus losing several electrons. Similar difficulties are met with in the cases of atoms with five or six electrons in the outer shell as in the case of seven electrons. The cases in which four or six electrons have been removed are, however, considerably simpler, since the electrons may then be given positions on a spherical surface such that the force acting on an electron is not dependent on the electron taken. In general, the atoms ionized to so great an extent will be comparatively few. In all our cases of ionized atoms, it must be remembered that it is not only the normal atom that is ionized; an atom may have given out several spectral lines before it becomes ionized. So a

large number of lines will be obtained depending upon the state of the atom when one or more electrons are removed from it. There is another form of ionization which is worthy of consideration. That is the case in which an electron has succeeded in penetrating the atom and reached the nucleus, thus temporarily reducing the positive charge and therefore giving a negative ion.

The fundamental mathematical expressions for such a case are found by extending our formulæ for the neutral atom for the case where the charge on the nucleus is 'n' instead of ten.

The angular velocities about the several axes will remain unchanged; and so only two equations will be required to determine the new values assumed by 'c' and 'r,' the radii of the outer and inner shells respectively.

We have three available equations; and expressing them in the notation previously used, we obtain:

$$\frac{(4n-1)e^2l}{4c^3} = \frac{e^2}{4l^2} + \frac{e^2l}{2s^3} + \frac{e^2(r+l)}{[(r+l)^2+s^2]^{3/2}} - \frac{e^2(r-l)}{[(r-l)^2+s^2]^{3/2}} + 2m\omega_1^2 l, \quad . \quad . \quad . \quad . \quad (\text{I}\,a.)$$

$$\frac{(4n-1)e^2l}{4c^3} = \frac{e^2}{4l^2} + \frac{e^2l}{2s^3} + \frac{e^2l}{[(r+l)^2+s^2]^{3/2}} + \frac{e^2l}{[(r-l)^2+s^2]^{3/2}} + m\omega_1^2 l + m\omega_3^2 l, \quad . \quad . \quad (\text{II}\,a.)$$

$$\frac{ne^2}{r^2} = \frac{4e^2(r+l)}{[(r+l)^2+s^2]^{3/2}} + \frac{4e^2(r-l)}{[(r-l)^2+s^2]^{3/2}} + \frac{e^2}{4r^2} + 2m\omega_1^2 r, \quad . \quad . \quad . \quad . \quad (\text{III}\,a.)$$

which are obtained from equations (I.), (II.), and (IV.), replacing the nucleic charge of '$10e$' by 'ne,' and remembering that $\omega_1 = \omega_2$.

The result of the alteration will be that the electrons will move further out from the nucleus, since 'n' is of necessity less than ten. In consequence of this, the frequency of the spectral lines produced by such ionized atoms will be of a lower frequency than those produced by the neutral atom. There will in consequence be a larger number of lines in the part of the spectrum of greater wave-length. It is for this reason that there are so many lines in the orange, and red in the case of neon.

Feb. 13, 1922.

XV. *Absorption of Hydrogen by Elements in the Electric Discharge Tube.* *By* F. H. Newman, *Ph.D., F. Inst. P., A.R.C.S., Head of the Physics Department, University College, Exeter* *.

1. *Introduction.*

THE phenomenon of the disappearance of gas in the electric discharge-tube, and in the presence of incandescent filaments, has received much attention recently owing to its importance in technical applications. Langmuir † has shown that hydrogen disappears from a vacuum tube in which a tungsten filament is heated above 1000° C. This fact has been utilized by him in the removal of the last traces of gas in valves, and the effect has been termed a "cleaning up" one. The pressures at which he worked were very low; for example, he found that the pressure in a tube was lowered to 0·00002 mm. of Hg. Other gases, including nitrogen and carbon monoxide, are removed in a similar manner, and molybdenum, when incandescent, has the same effect as tungsten. In all cases Langmuir found that the cooling of part of the apparatus by means of an enclosure at liquid-air temperature greatly accelerated the rate of disappearance of the gases. In addition he noted an electro-chemical "clean up," which occurred at much lower temperatures of the filament, when potentials of over 40 volts were used in a way that caused a perceptible discharge through the gas.

More recently Campbell, conducting work for the General Electric Company ‡ and using incandescent filament cathodes in electric discharge-tubes, has made an exhaustive study of the "clean up" effect, and has come to the conclusion that there is much evidence for believing there exists an electrical action which is quite independent of the thermal action, and, providing the temperature of the filament is kept below that at which the chemical "clean up" occurs, the effect appears to be one dependent only on the electrical discharge. In the case of the disappearance of carbon monoxide there is proof of the conversion of this gas into carbon dioxide, and the action takes place more rapidly when part of the apparatus is cooled to liquid-air temperature. This has the effect of removing the carbon dioxide by condensation as quickly

* Communicated by the Author.

† Am. Chem. Soc. Journ. vol. xxxvii. (1915).

‡ Phil. Mag. vol. xl. (1920); vol. xli. (1921) & vol. xlii. (1921).

as it is formed. The presence of phosphorus vapour accelerates the rate of disappearance of all gases except the inert ones, and much lower final pressures are attained. This, the author believes, is due to the deposition of the gas on the walls of the vessel, this deposit then being covered with a layer of red phosphorus formed by the electric discharge passing through the phosphorus vapour. The covering of red phosphorus prevents liberation of the hydrogen by bombardment of the ions, and at the same time provides a new surface on which further gas can be deposited.

The problem of the disappearance of the gas is a very complicated one, owing to the many factors to be considered. The walls of the vessel and the electrodes will certainly receive some of the gas, although the latter may not disappear in its original state. There will be chemical changes occurring in the volume of the gas, such as the conversion of carbon monoxide into carbon dioxide, and, in addition, any other elements present in the discharge-tube, either in the form of vapour or solid, will affect materially the rate of disappearance of the gas and the final pressure reached.

The author * has shown previously that various substances present on the electrodes of a discharge-tube alter considerably the amount of gas that can be caused to disappear when an electric discharge is passing. In particular, phosphorus, sulphur, and iodine cause both hydrogen and nitrogen to be absorbed at a very great rate, and a high vacuum is quickly produced as a result. This action of phosphorus has been used for many years to obtain and maintain very low pressures in valves. These three elements stand out as being very effective even at high pressures, but other substances which were tested in a similar manner did not appear to absorb hydrogen. On the contrary, gas appeared to be liberated. This effect can be explained as follows. At pressures above 1 mm. of Hg. a certain amount of the gas in a discharge-tube becomes occluded within the walls. This gas will be liberated when the walls are bombarded by the ions produced by an electric discharge. This effect will mask any disappearance. If, however, the tube is heated almost to the softening point of glass and highly exhausted, then on admitting hydrogen at a small pressure such as 0·1 mm. of Hg., very little occlusion of the gas within the walls will take place, and on passing the electric

* Newman, Proc. Roy. Soc., A. vol. xc. (1914); Proc. Phys. Soc. vol. xxxii. (1920) & vol. xxxiii. (1921).

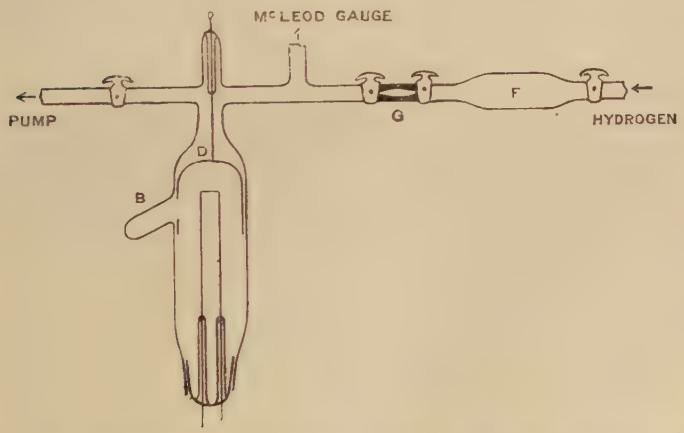
discharge practically no hydrogen will be liberated from the walls by bombardment with the ions. If there is any absorption of the gas, this effect will not be masked by the liberation of the gas from the walls or electrodes.

The object of the present work was the study of the behaviour of hydrogen in the presence of various elements in a discharge-tube when a current was passing through it. The pressures of the gas in these experiments were much lower than those used by the author in the papers quoted above, but they were much greater than those used by the previous investigators—Langmuir and the General Electric Campany.

2. *Description of Apparatus.*

At gas-pressures below 0·1 mm. Hg. it is difficult to obtain a current through a discharge-tube unless very high potentials are used. A valve also must be placed in the circuit to make the discharge unidirectional. This entails further diminution

Fig. 1.



of the current. By using a Wehnelt cathode the potential required was greatly reduced. The apparatus employed is shown in fig. 1. The incandescent filament was a strip of platinum foil 5 mm. long and 3 mm. wide. As the discharge-tube had to be thoroughly cleaned after each experiment, the cathode was sealed in a glass stopper which could be removed when the tube was cleaned. This necessitated the use of tap-grease, but the vapour arising from it did not appear to affect the results at the pressures used.

Previous experiments had shown that elements such as sodium and potassium are only effective in causing the disappearance of gas in the electric discharge-tube if the surface of the element is clean, and if it has been prepared in vacuum. Accordingly the substance under test was placed on the platinum foil forming the cathode, and after the tube had been heated almost to the softening point of glass and exhausted, the element on the foil was vaporized by passing an electric current through the latter. In this way the substance was then deposited on the inner surface of the anode D and an uncontaminated surface obtained. The anode was of aluminium and was cylindrical in shape, fitting very closely to the glass walls. Enclosing the cathode in this way, the effect of the surface of the glass on the absorption of the gas was minimized. A side tube B was used to contain the easily volatile elements such as phosphorus, sulphur, and iodine. An aperture was made in the anode opposite the mouth of B so that the vapour of the substance from B could pass through and be deposited on the inner surface of D. The pressures of the gas were measured with a McLeod gauge. The hydrogen was prepared by the electrolysis of barium hydroxide and stored in a reservoir. This method of preparing the gas ensues great purity. Any oxygen present was removed by passing the gas through a bulb containing sodium-potassium alloy. Phosphorus pentoxide in F removed any water-vapour, and of course the alloy was effective in this respect also. The gas could be admitted to G, which was a known volume (0·051 c.c.) enclosed between two taps. A definite volume of gas at a known pressure could thus be admitted to the discharge-tube. From observations of the pressure in the tube before and after a discharge had passed, the actual volume of gas — at atmospheric pressure — which had disappeared could be calculated.

The current through the discharge-tube was kept constant by altering the filament current, and was measured with a galvanometer. In previous experiments the quantity of electricity passing through the tube while absorption was taking place had been measured with a water voltameter, but in the present work this method was not sensitive enough.

After deposition of the substance on the anode D, the tube was again highly exhausted to remove any gases liberated from the volatized substance. The tube was placed in an enclosure maintained at −40° C. while absorption of gas was in progress.

Observations were then taken of the changes in pressure due to the disappearance of the hydrogen when an electric discharge passed through the gas. The results obtained are shown in the accompanying table.

The accelerating potential was 94 volts, obtained by using small accumulators. The current through the tube was kept constant, and was 546 micro-amps.

Each set of readings corresponds to an electric discharge for ten minutes, except in the cases of sulphur, phosphorus, and iodine, where the observations were taken at intervals of two minutes—*i. e.*, with sodium the pressure changed from 743×10^{-3} mm. of Hg. to 336×10^{-3} mm. of Hg. in ten minutes, while with sulphur the pressure was lowered from 740×10^{-3} mm. of Hg. to 329×10^{-3} mm. of Hg. in two minutes.

The amount of hydrogen which would be liberated from a water voltameter in 10 mins. by the same current is 39×10^{-3} c.c. at atmospheric pressure.

As the gas may disappear into the walls of the anode even in the absence of any substance on the anode, and as the glowing filament may affect the rate of disappearance, preliminary observations were always made when an electric discharge passed through the tube without the substance present on the anode. The volume of gas which disappeared owing to these two effects was always very small compared with that which was absorbed when the element under test was on the anode.

3. *Experimental Results.*

After each element had been tested, the tube was heated to 300° C., and the volume of gas reliberated was calculated from the observed change of pressure. The amount thus recovered varied considerably, but was always less than that which had disappeared. This evolved gas was again absorbed when a discharge was passed, and it is evidently hydrogen in the same condition as it was before disappearance.

If, after the gas had disappeared, a fresh amount of hydrogen was admitted, the volume which disappeared on discharge was reduced. For example, with sodium and the gas pressure at 743×10^{-3} mm. of Hg., the vacuum was reduced to 96×10^{-3} mm. Hg. before the action ceased. Admitting a further supply of gas to the tube, the pressure fell from 743×10^{-3} mm. Hg. to 233×10^{-3} mm. Hg. Repeating the process again, the pressure fell from 743×10^{-3} mm. Hg. to 436×10^{-3} mm. Hg., and then the

Element.	Initial Gas Pressure. mm. Hg. $\times 10^{-3}$	Final Gas Pressure. mm. Hg. $\times 10^{-3}$.	Volume of Gas absorbed. c.c. $\times 10^{-3}$.	Pressure of Gas at which action ceased. mm. Hg. $\times 10^{-3}$.
Sodium	743	336	28	96
	336	123	14	
Potassium	726	349	26	84
	349	163	13	
Sodium-Potassium Alloy.	738	392	23	110
	392	188	15	
Sulphur	740	329	28	26
	329	145	13	
	145	58	6	
Phosphorus	748	352	27	14
	352	147	14	
	200	44	7	
Iodine	763	338	28	124
	338	150	13	
Arsenic	750	394	24	108
	394	206	13	
Cadmium	744	463	19	284
	463	321	10	
Calcium	731	404	23	152
	404	308	7	
Tin	750	386	25	131
	386	264	8	
Zinc	753	431	22	276
	431	306	8	
Thallium.....	732	560	12	297
	560	399	11	
Mercury, Lead	Hydrogen was liberated and not absorbed.			

absorption ceased. There appears to be a fatigue effect whereby the actual amount of gas which can be absorbed by any surface is limited. This fatigue effect may be due to three causes. If the disappearance of the gas depends on chemical action, the latter will occur mainly at the surface of the element. The formation of a chemical compound will thus protect the rest of the substance from the action, and the process will gradually cease. If, on the other hand, the effect is due to a deposition of the gaseous atoms on

the surface, as Langmuir suggests, these atoms will diffuse slowly into the substance. The atoms arriving later will have less area on which deposition can take place. A limit to the action will be reached when the number of atoms deposited is equal to the number set free by the bombardment of the surface by the ions.

After absorption, the proportion of the hydrogen reliberated when the tube was heated to 300° C. varied considerably in different cases, not only with different elements, but also with the same element. This is to be expected when it is remembered that the thickness of the substance deposited on the anode varied with different substances.

The accelerating potential affected to some extent the rate of disappearance of the gas and also the final pressure attained. Owing to liberation of the gas by the bombardment with the ions, the final pressure reached must depend on this reverse action, and the greater the accumulation of the gas on the surface of the anode, the greater will be the amount of gas evolved.

With sodium on the anode a potential of 94 volts reduced the gas-pressure from 743×10^{-3} mm. Hg. to 123×10^{-3} mm. Hg. in the course of 20 minutes. When the potential was lowered to 54 volts, the pressure fell from 743×10^{-3} mm. Hg. to 476×10^{-3} mm. Hg. in the same time-interval. The final pressures reached before absorption ceased were 96×10^{-3} mm. Hg. and 202×10^{-3} mm. Hg. respectively. The current through the discharge-tube was kept constant throughout.

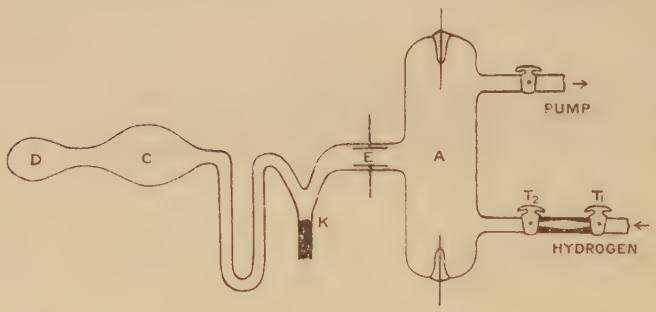
The principles of the disappearance of the gas will be discussed later, but there are certain features of the phenomenon which can be traced to chemical actions.

Many of the elements tested combine with hydrogen at high temperatures to form chemical compounds which are very stable. Any chemical action occurring in the present experiments cannot be due to the heat, as the discharge-tube was maintained at −40° C., and the incandescent filament was always at a lower temperature than that at which Langmuir found chemical action occurred with hydrogen. The effect may be caused by "activation" of the gas, the latter assuming some modification under the action of the electric discharge. In the above experiments the amounts of gas absorbed were so small that it would be extremely difficult to detect the existence of any chemical compound in the tube. In order to increase the amount of gas absorbed,

and test for any chemical compounds formed, a modified form of the apparatus was used, as shown in fig. 2.

Pure hydrogen could be admitted to the discharge-tube A in small volumes by manipulation of the taps T_1 and T_2. Two strips of platinum foil, about 10 cms. long, were sealed in the tube E. These strips fitted closely to the glass surface. A potential difference of 600 volts was applied between these strips by means of small accumulators. In this way the ions actually present in E were removed while the discharge was proceeding in A. The tube E communicated with a mercury cut-off K, and a U tube immersed in an enclosure maintained at $-40°$ C. Sodium-potassium alloy was prepared in D, and after the whole of the apparatus had been evacuated, the alloy was run into C. In this way a bright and clean surface was obtained on the alloy in C. A

Fig. 2.



small volume of hydrogen was then admitted to A, the mercury cut-off preventing the gas from entering C and D. While the electric discharge was passing in A, the hydrogen was allowed to enter C by manipulation of the cut-off. Admitting successive volumes of hydrogen into A in this way, and each time allowing communication with C while the discharge was proceeding, an increasing amount of active gas entered C, and an effect was observed on the surface of the alloy. At first it appeared to be covered with a thin white crystalline compound when observed through a microscope. This white layer slowly changed, on the admission of more active gas, to a dark grey-coloured deposit. To show that this surface effect was not due to impurities in the hydrogen, previous experiments were made, the gas being admitted to C without the electric discharge proceeding. There were no surface effects then,

so it was concluded that some of the hydrogen assumes an active modification under the action of the electric discharge, and in this form it is able to form chemical compounds with the sodium and the potassium present in the alloy. The U tube in the enclosure at $-40°$ C. excluded the possibility of the action being due to the heat from the discharge-tube. The white crystalline compound which first appears is a mixture of the hydrides of sodium and potassium. The exact nature of the greyish-coloured product formed afterwards is unknown, but it is probably a solution of the hydrides in the alloy.

Water is evolved by an electric discharge when passed through any vacuum vessel. It comes from the glass, and would not be kept back by the trap cooled to $-40°$ C.; for at that temperature water substance has a vapour-pressure of about 0·1 mm. Hg. The presence of water-vapour "fouls" the surface of the alloy, but this fouling gives a black deposit on the surface which is quite different from the white crystalline layer observed in the present experiments. The black deposit consists of sub-oxides of sodium and potassium, and its appearance has been noted previously by the author *, although in that paper it was attributed to the hydrides. It has now been proved by chemical analysis that this black deposit does consist of the sub-oxides.

Sulphur was tested in the following manner:—A small piece of filter-paper, soaked in lead-acetate solution, was placed together with a small amount of the solution in D. The rest of the apparatus was separated from D by a mercury cut-off not shown in the figure. C contained sulphur which had been deposited in a thin film over the interior. After exhausting the whole of the apparatus to a pressure of about 5 mm. of Hg., the mercury cut-off between C and D was closed and the rest of the apparatus highly exhausted. Hydrogen was then admitted to A until the pressure was about 7 mm. Hg. While the electric discharge was passing, the mercury cut-off was opened. This was repeated many times, the pressure of the gas in A being gradually increased. Each time communication with D was established, any gaseous product formed in C was admitted to D. In the course of a few minutes the paper soaked with the lead-acetate solution turned black, showing the presence of a sulphide of hydrogen. This chemical compound must have been produced by the action of an active form of hydrogen on the sulphur. The surface of the mercury at the cut-off

* Proc. Roy. Soc. A. vol. xc. (1914).

also lost its bright appearance. This was due to the action on it of the sulphide of hydrogen. The mercury surface remained quite clear when the sodium-potassium alloy was tested.

Sulphur and the alloy were selected for tests because the chemical actions in these cases give rise to compounds whose effects can be noted easily. It is extremely difficult to examine phosphorus and iodine in this way owing to their high vapour-pressures. A possible test would be the comparison of the vapour-pressures before and after absorption of hydrogen had taken place.

These two experiments indicate that the chemical action is not due directly to the ions in the discharge-tube, as they were all eliminated by the charged platinum strips before reaching either the alloy or the sulphur.

Wendt * showed that hydrogen can be activated by the passage of α-rays through the gas, and it has been shown by the author † that the active modification so produced is able to react chemically with sulphur and the alloy of sodium and potassium.

4. *Discussion of Results.*

The disappearance of gas in a vacuum-tube is due probably to several principles, some of which may be fundamental. It is certain, however, that any attempt to explain the principles by the same theory would lead to conflicting results, but the processes occurring can be divided into two classes, chemical and mechanical. There is much evidence that the gas can be caused by the electric discharge to adhere to the solid parts of the discharge-tube in some manner which is at present unknown. In many cases a portion of the gas can be reliberated by heating the vessel, but no reason can be advanced for the non-reliberation of the whole of the gas which has disappeared.

Langmuir assumes in the paper previously quoted that the hydrogen in the presence of an incandescent filament undergoes dissociation. The gas shows abnormal thermal conductivity at high temperatures, due to its atomic nature. The dissociation does not occur apparently in the space round the wire, and is not due to the impacts of the gas molecules against its surface, but takes place only among the hydrogen molecules which have been absorbed by the metal of the wire. Some of the atoms leaving the wire do not meet other atoms,

* Nat. Acad. Sci. Proc. vol. v. (1919).
† Phil. Mag. vol. xliii. (1922).

owing to the low pressure, but diffuse into the tube cooled by liquid air, or become absorbed by the glass, and thus remain in the atomic condition. They retain all the chemical activity of the atoms. Langmuir also found that when the liquid air was removed, some of the atoms would come off the glass and recombine with other atoms to form molecules. These molecules could not be recondensed by replacing the liquid air. This gas which would not again disappear he termed a "non-recondensible" gas.

This hypothesis, which is applicable to very low pressures, cannot hold at the pressures used in the present work. The gas in the atomic condition can scarcely move from the discharge-tube for a considerable distance and still retain its atomic nature. The "non-recondensible" gas found by Langmuir is probably hydrogen in its normal state.

When nitrogen gas disappears in the discharge-tube, practically none of it can be reliberated, even when the tube is heated to the softening point. This fact indicates a striking difference between the disappearance of hydrogen and nitrogen. If chemical compounds are formed by the absorption of the gases, this difference can be explained in terms of the difference in the stability of the hydrides and nitrides produced.

The chemical action may take place between hydrogen and the vapour of the element, and also it may occur at the surface of the solid. The majority of the elements studied have such small vapour-pressures that a very small proportion of the action is due to the vapour. The active condition of the gas must be caused by the ions, although results seem to indicate that the number of active atoms or molecules in the gas is of a much higher order than the number of ions present in the gas at the instant of recombination.

The absorption is not due entirely to chemical action, as the law of constant proportions does not seem to be followed. It is of significance, however, that the rate of disappearance of the gas increases, and the final pressure attained decreases, as the temperature of the discharge vessel is lowered. This arises from the lowering of the vapour-pressure of the compounds produced, with the result that the final pressure reached is lowered.

Although the formation of hydrogen sulphide in the discharge-tube by the action of the activated hydrogen on the sulphur will not explain the disappearance of the gas, it does indicate the production of a modified form of the gas which is able, possibly, to form other compounds with sulphur in addition to hydrogen sulphide.

That the mechanical deposition of gas on the walls of the discharge vessel will not account entirely for the disappearance of the gas is shown by the difference in the behaviour of nitrogen and hydrogen with phosphorus, sulphur, and iodine. Practically none of the nitrogen can be reliberated by heating, but a large proportion of the hydrogen is evolved.

There is reason for believing that the modification of hydrogen is triatomic in nature. Wendt has shown in the paper previously quoted that hydrogen drawn from a tube through which an electric discharge is passing contains a small quantity of H_3. Probably monatomic hydrogen is first formed, and owing to collision with neutral molecules of the gas, H_3 then appears. The monatomic gas may be produced originally by the action of the swift-moving electrons on the molecules. Wendt and Grubb * have also shown that N_3 is produced when an electric discharge passes through nitrogen. Thomson † found evidence of H_3 in his positive-ray experiments. It is this triatomic form of hydrogen which is effective in the production of chemical compounds in the electric discharge-tube.

XVI. *Molecular Thermodynamics.* II. *By* Bernard A. M. Cavanagh, *B.A., Balliol College, Oxford* ‡.

I. Molecules, Thermodynamics, and Quantum Theory.

IN developing a molecular treatment of the thermodynamics of dilute solutions in simple solvents, Planck § determined the form of the integration constants in the entropy function by a method which was at the time the subject of some controversy.

M. Cantor ‖ objected that the hypothetical transition to the gaseous state without change of the molecular composition was not even *theoretically* possible, "since there probably existed in the liquid state, complex molecules whose existence was inseparably connected with the condensed state of the phase, and entirely incompatible with a state of high temperature and low pressure.

* Science, vol. lii. (1920).

† Proc. Roy. Soc. A. vol. lxxxix. (1913).

‡ Communicated by Dr. J. W. Nicholson, F.R.S.

§ 'Thermodynamics,' 1917 (Trans. Ogg), pp. 225-226. *Or see* Phil. Mag. xliii. p. 608 (1922).

‖ *Ann. der Phys.* x. p. 205 (1903).

In reply, Planck * pointed out that the *theoretical* possibility of the ideal transition depended only on the fact that the numbers of the various molecular species were, together with temperature and pressure, the independent variables which determined the phase.

Now the present author would suggest that Planck's reply can be construed (and, to be unanswerable, *must* be construed) as a rider to the definition of the terms "molecule" and "chemical compound," for the purposes of molecular thermodynamics, viz. :—"That the numbers of the various molecular species *can* be considered, *together with temperature and pressure*, as the independent variables determining the phase," or, in other words, "that it *shall* be theoretically sound to conceive any desired change of temperature and pressure of the system as taking place without change in the numbers present of the various molecular species †."

Only with this rider to our definition can it be laid down, for instance, that the "mass-action" equilibrium law must be obeyed in sufficiently dilute solution, for it is to be observed that purely "general" thermodynamics has no cognisance of molecules, but takes for its independent variables, besides temperature and pressure, the masses of the "components." [See next section of this paper.]

The misconceptions which have so long stood in the way of a satisfactory general theory of electrolytic or "conducting" solutions seem sufficiently to illustrate the indispensability of this postulate.

A parallel illustrating its significance may be drawn from the dynamical theory of chemical combination and dissociation.

The classical dynamical conception of a binary molecule (for example) was a pair of simpler molecules (or atoms) moving relatively to one another in closed orbits, and the principle of the conservation of energy forbade the spontaneous dissociation of such a "molecule," requiring that its disruption should depend on collision with another molecule.

The well-known fact that dissociation is (at constant temperature) independent of collision-frequency, showed the inadequacy of this conception, and pointed to some property, in the "forces" producing and maintaining a molecule, altogether incompatible with the older or "continuous" dynamics.

* *Ann. der Phys.* x. p. 436 (1903).

† Intermediate and final states being unstable, of course, in general.

The same difficulty arose when a dynamical explanation of the law of mass-action was attempted, the essential continuity of action of "physical" forces * standing in the way, and Boltzmann had to assume—with conscious artificiality—discontinuity in a field of force in order to arrive at the desired result.

It seems indeed that, besides the more obvious and less peculiar properties of shortness of range, "specificity," and saturability, there is a quality of *discontinuity of action* (in time or in space or in both) which distinguishes "chemical forces" from "physical," the distinction being sharp so far as we can yet see.

We find then an absence of direct dependence of dissociation upon the *thermal motion* closely connected with the possibility of accounting for the law of mass-action dynamically.

The parallel with our "rider" and its indispensability as a basis for the deductions of molecular thermodynamics is significant.

The transitory orbital system which was the older "physical" conception of a molecule, and which quite probably occurs in all dense gases and liquids †, is clearly quite directly dependent on the thermal motion, is, in fact, itself merely an "episode" in that motion, and cannot in any sense be regarded as fulfilling the requirement of our rider. We cannot, therefore, predict from thermodynamics the "mass-action" equilibrium law for the "reactions" of such "molecules" under any circumstances, for we cannot treat them as molecules for the purposes of molecular thermodynamics. And, in parallel, we find that dynamical theory is unable to predict the law of mass-action for such "molecules."

The electrolyte question provides an important application of these considerations.

It has frequently been supposed that a pair of ions, closely linked by their electrostatic fields alone, must be regarded as a molecule, and should behave thermodynamically as such. Electrostatic forces as we know them, however, are typical

* That is, "forces" within the conception of the older physics. It seems convenient to use, in contradistinction, the term "chemical forces" for the "forces" or means by which a molecule is formed or "bound" [see end of this section] and held together, and the expression may find some further justification in the fact that in the nature of these latter "forces" lies, probably, the key to all the facts and phenomena of chemistry.

† Compare here Cantor's objection, mentioned above.

"physical" forces, and it is our present conclusion that such forces are capable of forming only transitory associations—"episodes in the thermal motion"—essentially different from what we regard as molecules.

When, as in the weaker acids and bases almost certainly, we really have partial ionization, or rather partial association of the ions to form "undissociated molecules," the latter must be regarded as produced and held together not by such ordinary electrostatic forces, but by "chemical forces" with the peculiar property already discussed.

With regard to "strong electrolytes," the work of Debye and Bragg gives good reason to believe that the molecule in the salt crystal is *the ion*. If this be so, it appears necessary to admit that *the solid salt is essentially a mixed crystal* whose special simplicity and homogeneity is due simply to the polarity of the electrostatic forces which dominate its "growth."

Now, it would seem altogether inconsistent to suppose that the chemical "association" which does not take place in the intimate contact of the solid state, ensues when the ions are dispersed in a solvent, so that until the calculated effect of the electrostatic forces between the ions upon their thermodynamic behaviour can be shown to be inadequate when compared with experiment, the "complete-ionization" theory seems the only rational theory for strong electrolytes.

Another application of the above general conclusions is to be found in the important question of "solvation" of solutes, which is treated in a paper to follow this.

There occur in the literature of this subject such statements or suggestions as that "the solvates need not be definite chemical compounds," and vague theories of the "solvate molecule" as a mere indefinite conglomerate. From the preceding, at any rate, it is our conclusion that, unless the "solvate molecule" is produced and maintained by "chemical forces" in the sense already considered, so that it fulfils the requirement of our "rider," it will not, for the purposes of molecular thermodynamics, be a molecule at all.

With regard to this remarkable characteristic of "chemical forces" which appears to be reduced to its lowest terms in the expression "discontinuity of action," this seems to mark out the problem of molecule formation (including, be it noted, reaction-velocity) as one of those many whose solution may be hoped for from the new quantum-dynamics of phenomena on the atomic scale. Indeed, it is tempting to believe that Bohr's conception of "electron-binding" may be the solution

in embryo of the larger and more complex problem of "atom-binding," and that in his distinction between "bound" and "unbound" electrons in the atom, we may have in its simplest aspect the distinction between "chemical" and "physical" forces.

II. Molecular Thermodynamics.

In general thermodynamics, which is based on, and applies to experience, the independent variables, besides temperature and pressure, are the masses of the "components," and these are reduced to the minimum necessary to define the system under all circumstances not conventionally or practically excluded from consideration.

To take a familiar example, hydrogen and oxygen will suffice as the components of a system containing in addition water, provided low temperatures are excluded from consideration, or the presence of efficient catalysts is assumed. In so far as we may suppose that the decomposition and formation of water do proceed even at low temperatures in the absence of catalysts, though at an immeasurably small rate, it is clear that theoretically the two components would always suffice for this system *if sufficient time were allowed.*

And conversely *they would never suffice if the rate of experimentation were sufficiently increased.*

Striking examples of the practical reality of this *entry of the time factor* into the question of the necessary number of components, have been given in recent years by the work of A. Smith and of A. Smits, who by increasing the rate of experimentation, have increased the number of components necessary to describe certain systems, the latter author having propounded an interesting theory of allotropy on the basis of his experiments.

Now we can conceive this carried far beyond the bounds of purely practical limitations, and the question arises, "How far?"

The atomic or elementary theory of matter is introduced when we say that at one extreme, when unlimited time is available, the elementary atomic species will be necessary as well as sufficient as the components of any system.

Starting from this extreme and increasing the rate of experimentation we can imagine one complex after another of these elementary atoms (as its rate of formation and decomposition ceases to be great in comparison with the rate of experimentation) taking its place in the list of "components necessary to describe the system." Remembering,

however, that the masses of the components are *together with temperature and pressure* the independent variables, we see that there may be a limit, for it must be *theoretically* possible to alter temperature and pressure so quickly that the numbers present of these complexes which we are admitting as components remain sensibly unaltered *during the change*. According to the view put forward in the previous section, it involves us in a definite postulate bearing upon the nature of molecules and of chemical change, when we say that the limit is reached *when and only when* every molecular species which can be formed in the system has taken its place in the list of "components."

Proceeding in theory to this limit, we obtain the general expression for ψ which is referred to in the sequel as the "molecular expression for ψ," and as Planck showed, we can determine it completely when it is linear by "connecting-up" with the known properties of the low-pressure gaseous mixture.

When the expression is not linear, the higher or "general" terms are subject only to a single limitation inherent in Planck's method, as pointed out in the previous paper *. Observing that the "general" terms in the corresponding expression for U † do not involve "chemical" energy, the present author also suggested and illustrated ‡ the interesting possibility of employing ordinary dynamical theory, at least as a valuable aid, in determining and interpreting the form of these "general" terms.

This "molecular expression for ψ," however, will clearly not in general correspond with our experiments carried out under ordinary conditions. They will correspond with an expression of the "general thermodynamic" type in which the components are appropriate to the conditions of experiment.

The theoretical problem then presents itself of connecting this "experimental" expression for ψ, in a manner at once rigorous—that is trustworthy—and practically effective, with the "molecular" expression and its possibilities of theoretical interpretation.

The treatment of two important problems of this kind has been attempted.

Planck pointed out that when a single molecular species

* Phil. Mag. xliii. p. 606 (1922).

† Entirely analogous considerations apply to V, but ordinarily owing to the low pressures used, V figures relatively negligibly in the determination of ψ [*cf.* footnote, p. 630, Phil. Mag. xliii. (1922)].

‡ *Loc. cit.* p. 625.

greatly preponderates, it is a matter of mathematical necessity that the "molecular expression for ψ" should take (in the limit) a linear form, and it was to this type of "dilute solution" that Planck confined himself, arriving readily at the Raoult-van't Hoff "laws of dilute solution." Dilute solutions in a liquid paraffin would be of this type.

Van Laar * took the linear expression as the criterion of "perfect solution" in general, and not making the approximations which Planck, considering very low concentrations, had made, was able to show that the Raoult-van't Hoff laws formed too restricted a criterion when the solution was very dilute.

He, however, considered only solutions in which the solvent was of the same type as that of Planck, viz.: a single molecular species.

In view of the fact that the Raoult-van't Hoff laws have been found to hold for dilute solution in our common and useful solvents, which are certainly not of the type considered by Planck and Van Laar, the present author was led to the problem of "complex solvents," which will be the first illustration of the theoretical problem outlined above. A preliminary treatment appeared in the first of these papers, but a more complete and rigorous treatment is now presented.

The second illustration will be the problem of partially "solvated" solutes, a discussion of which will follow that of "complex solvents."

The first result is that the Raoult-van't Hoff laws have been rigorously predicted for extreme dilutions in such solutions. It is shown, in fact, how the "experimental" expression for ψ simulates, in the limit, the "molecular" expression in form.

But further the way is prepared for the thorough investigation of middle and high concentrations in such solutions. To this end the "linear" terms in the experimental expression for ψ, which simulate and replace the simple linear terms in the molecular expression, have been treated with some thoroughness and rigour, these being the terms to which the expression reduces when the solution is "perfect."

When the quite practical criteria thus provided are applied, the belief that "perfect solution" always ceases in these "complex" solutions at quite low concentrations may be largely dispelled.

In simple solutions of the kind considered by Planck and

* *Z. f. Phys. Chem.*, several papers, 1903 etc.

Van Laar we find "perfect solution" persisting up to very high concentrations—sometimes over the whole range.

The "general" terms of which (excepting the case of electrolytes) little or nothing is yet known, have been touched on only in so far as the treatment of the "linear" terms involves (in general) a certain very slight alteration in the division into "linear" and "general" terms, which may sometimes have to be taken into account when dynamical theory is employed.

Experimental determination of the "general" terms for comparison with theory must of course be preceded by knowledge of the "linear" terms,—hence again the need for rigorous and thorough treatment of the latter.

III. Complex Solvents.

The importance of this question of "complex" (polymerized and mixed*) solvents is sufficiently obvious when it is considered that of this type are most of our best solvents, probably all our "ionizing" solvents, and, chief of all, water.

We have to consider a solution consisting of the solute-molecular species, m_1, m_2, ..., m_s, ..., in addition to the various species m_{01}, m_{02}, ..., which constitute the solvent.

Concentrations being expressed in gram-molecules per gram of solvent, the "molecular" expression for ψ is

$$\psi = \Sigma n_{01}\left(\phi_{01} - \mathrm{R}\log\frac{c_{01}}{\Sigma c_{01} + \Sigma c_s}\right)$$

$$+ \Sigma n_s\left(\phi_s - \mathrm{R}\log\frac{c_s}{\Sigma c_{01} + \Sigma c_s}\right) + \mathrm{R}M_0\Sigma\phi_x' f_x(c_{01}\ldots c_s\ldots), \quad (1)$$

which is of the form

$$\psi = \Sigma n_{01}\frac{\partial\psi}{\partial n_{01}} + \Sigma n_s\left(\frac{\partial\psi}{\partial n_s}\right)_{n_{01}}\dagger, \quad . \quad . \quad . \quad (2)$$

the several solvent-molecular species appearing as separate

* *Mixed solvents*, while submitting to the same theoretical treatment as the merely polymerized solvents, present certain peculiar difficulties and some interesting possibilities with which it is hoped to deal at length in a later paper.

† A suffix outside a bracket is used to indicate, in less obvious cases, independent variables which are held constant in a partial differentiation,—a well-known usage. A single suffix may be used briefly for a whole series, as n_{01} here standing for $n_{01}, n_{02}, \ldots\ldots$

components, whereas the "experimental" expression must have the form

$$\psi = M_0 \frac{\partial \psi}{\partial M_0} + \Sigma n_s \left(\frac{\partial \psi}{\partial n_s}\right)_M \quad . \quad . \quad . \quad . \quad . \quad (3)$$

the solvent appearing as one component only.

Clearly

$$\left(\frac{\partial \psi}{\partial n_s}\right)_{M_0} = \left(\frac{\partial \psi}{\partial n_s}\right)_{n_{01}} + \left[\Sigma \frac{\partial \psi}{\partial n_{01}} \cdot \frac{\partial n_{01}}{\partial n_s}\right]_{M_0} , \quad . \quad (4)$$

but the relation we have to use in making the change of variables is that given by the condition for chemical equilibrium among the molecules of the solvent, viz. :—

$$\left[\Sigma \frac{\partial \psi}{\partial n_{01}} \cdot dn_{01}\right]_{M_0} = 0. \quad . \quad . \quad . \quad . \quad . \quad (5)$$

so that

$$\left(\frac{\partial \psi}{\partial n_s}\right)_{M_0} = \left(\frac{\partial \psi}{\partial n_s}\right)_{n_{01}} , \quad . \quad . \quad . \quad . \quad . \quad (6)$$

and comparing then (2) and (3),

$$\frac{\partial \psi}{\partial M_0} = \Sigma c_{01} \frac{\partial \psi}{\partial n_{01}} , \quad . \quad . \quad . \quad . \quad . \quad (7)$$

which could be regarded as physically obvious, as was done in the preliminary treatment (previous paper).

We shall abbreviate (1) by writing $\bar{m}$ for $\frac{1}{\Sigma c_{01}}$, C for Σc_s, and $M_0 G'$ for the "general" terms, so that,

$$\psi = \Sigma n_{01}\left(\phi_{01} - R \log \frac{\bar{m} c_{01}}{1 + \bar{m} C}\right) + \Sigma n_s \left(\phi_s - R \log \frac{\bar{m} c_s}{1 + \bar{m} C}\right) + M_0 G' . \quad (8)$$

It will be convenient also to write G_{01}' for $\frac{\partial}{\partial n_{01}}(M_0 G')$, the "general" terms in $\frac{\partial \psi}{\partial n_{01}}$; similarly G_s' for those in $\frac{\partial \psi}{\partial n_s}$, etc.

Now, in the first place, we have to show that it is permissible to assume that (8) has already been so arranged that G', G_s', G_{01}', etc., all vanish in the limit when C becomes

very small. This is essential to the rigour of the demonstration that the Raoult-van't Hoff laws are still the limiting laws of dilute solution when the solvent is complex.

Now G′ is, at constant temperature and pressure, a function of $c_{01}, c_{02}, \ldots\ldots$, as well as of $c_1, c_2, \ldots\ldots$ Owing to the chemical equilibrium controlling $c_{01}, c_{02}, \ldots\ldots$, however, these quantities have, in the pure solvent, values depending only on temperature and pressure, and the *departures* from these limiting values, caused by the presence of solutes, will clearly decrease with the concentrations of those solutes. Thus as C diminishes, the *ranges of variation to be considered* of the variables $c_{01}, c_{02}, \ldots\ldots$ are progressively limited, the same being obvious in the case of $c_1, c_2, \ldots\ldots$

But clearly any finite, continuous, differentiable function of several variables must behave as a linear function if the range of variation considered of every variable is sufficiently limited *.

Thus

$$\operatorname*{Lt}_{C\to 0} G' = \Sigma c_{01} l_{01} + \Sigma c_s l_s, \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

i. e.

$$\operatorname*{Lt}_{C\to 0} (M_0 G') = \Sigma n_{01} l_{01} + \Sigma n_s l_s, \quad . \quad . \quad . \quad . \quad (10)$$

where $l_{01}, \ldots\ldots, l_1, \ldots\ldots$, depend only on temperature and pressure, being, in fact, limiting values of $G_{01}', \ldots\ldots, G_1' \ldots\ldots$, respectively.

But clearly $l_{01}, \ldots\ldots, l_1, \ldots\ldots$, can be transferred to, and included in the "linear" terms, — in $\phi_{01}, \ldots\ldots, \phi_1$, respectively — whereupon the residual "general" terms will satisfy the requirement that G′, G_s', G_{01}', etc., should all vanish as C becomes very small. We shall assume that in (8) this adjustment has already been carried out.

Returning now to (7) and comparing with (8), we see that

$$\operatorname*{Lt}_{C\to 0} \frac{\partial \psi}{\partial M_0} = \operatorname*{Lt}_{C\to 0} \left[\Sigma c_{01} \frac{\partial \psi}{\partial n_{01}} \right]$$

$$= \operatorname*{Lt}_{C\to 0} \left[\Sigma c_{01} (\phi_{01} - R \log \overline{m}\, c_{01}) \right] = \phi_M \text{ (say)}, \quad . \quad (11)$$

since, in the pure solvent, $c_{01}, \ldots\ldots$, and $\overline{m}$ assume values dependent only on temperature and pressure.

* Merely the obvious property of tangency in n-dimensions. The theorem quoted by Planck in treating simple solvents ['Thermodynamics' (Trans. Ogg), 1917, p. 225] appears to be the particular case of this, when the ranges of variation of the variables are *all* located (as here in the case of $c_1, c_2, \ldots$.) close to zero.

Then we have

$$\frac{\partial \psi}{\partial M_0} = \phi_M + \int_{C=0}^{} d\left[\Sigma c_{01} \frac{\partial \psi}{\partial n_{01}}\right] : \quad . \quad . \quad . \quad (12)$$

but again remembering

$$\Sigma \frac{\partial \psi}{\partial n_{01}} dc_{01} = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

we get

$$\frac{\partial \psi}{\partial M_0} = \phi_M + \int_{C=0}^{} \Sigma c_{01} d\left(\frac{\partial \psi}{\partial n_{01}}\right) : \quad . \quad . \quad . \quad (14)$$

and so from (8)

$$\frac{\partial \psi}{\partial M_0} = \phi_M + \int_{C=0}^{} \Sigma c_{01} \left(-R d \log \frac{\overline{m}C_{01}}{1+\overline{m}C}\right) + \int_{C=0}^{} \Sigma c_{01} dG_{01}' \quad . \quad (15)$$

$$= \phi_M + R \int_{C=0}^{} \frac{1}{m} d \log (1 + \overline{m}C + \int_{C=0}^{} \Sigma c_{01} dG_{01}. \quad . \quad (16)$$

Now (6) with (8) gives

$$\left(\frac{\partial \psi}{\partial n_s}\right)_{M_0} \left[= \left(\frac{\partial \psi}{\partial n_s}\right)_{n_{01}} \right] = \phi_s - R \log \frac{\overline{m}C_s}{1+\overline{m}C} + G_s'. \quad (17)$$

And if we write (16) in the form

$$\frac{\partial \psi}{\partial M_0} = \phi_M + R \int_{C=0}^{} \frac{1}{m} d \log (1 + \overline{m}C) + G_M, \quad . \quad (18)$$

it is at once clear that the "general" terms thus adopted (and therefore, of course, the "linear" terms similarly) are connected by the Gibbs fundamental relation [see note at end of this paper], for $(M_0 G')$ being a function of $n_{01} \ldots\ldots n_1 \ldots\ldots$, homogeneous and of the first degree,

$$\Sigma n_{01} dG_{01}' + \Sigma n_s dG_s' = 0 ;$$

i. e.
$$M_0 \Sigma c_{01} dG_{01}' + \Sigma n_s dG_s' = 0,$$

i. e.
$$M_0 dG_M + \Sigma n_s dG_s' = 0^*, \quad . \quad . \quad . \quad (19)$$

which means that $[M_0 G_M + \Sigma n_s G_s']$ or (say) $M_0 G$ is, *as a*

* The Gibbs relation might indeed have been used to obtain (16) direct from (17), ϕ_M appearing as the integration constant. The above treatment [(12) to (16)] appeared, however, to be more interesting, and to introduce ϕ_M in a more natural and illuminating manner, in its relation to the original "molecular" expression for ψ.

function of M_0 *and* $n_1 n_2 \ldots\ldots$, homogeneous and of the first degree. And the same will hold for the "linear" terms, (if $(M_0 G)$ be accepted as the new "general terms"), a point of obvious importance since the "linear" terms must alone remain when the solution is "perfect." $\frac{\partial \psi}{\partial n_s}$ retains the simple form (17), which may now be written

$$\frac{\partial \psi}{\partial n_s} = \left(\phi_s - R \log \frac{\overline{m} c_s}{1 + \overline{m} C}\right) + G_s, \quad . \quad . \quad (20)$$

since

$$G_s = \left(\frac{\partial (M_0 G)}{\partial n_s}\right)_{M_0} = \frac{\partial G}{\partial c_s}, \quad . \quad . \quad . \quad . \quad (21)$$

(G being a function of $c_1 c_2 \ldots\ldots$ only (besides T and p). Also, of course,

$$G_s = G_s' = \left(\frac{\partial (M_0 G')}{\partial n_s}\right)_{n_{01}} \quad . \quad . \quad . \quad . \quad (22)$$

and

$$M_0 G = M_0 G_M + \Sigma n_s G_s. \quad . \quad . \quad . \quad . \quad (23)$$

We observe, however, that this convenient arrangement involves theoretically a certain definite (though probably always very small) change in the division into "linear" and "general" terms, since

$$M_0 G' = \Sigma n_{01} G_{01}' + n_s G_s';$$

i. e.

$$G' = \Sigma c_{01} G_{01}' + \Sigma c_s G_s$$

$$= G + \int_{C=0} \Sigma G_{01}' dc_{01}. \quad . \quad . \quad . \quad . \quad . \quad (24)$$

Now it is readily shown that

$$G_{01}' = \frac{\partial G'}{\partial c_{01}} + m_{01}\left[G' - \Sigma c_{01} \frac{\partial G'}{\partial c_{01}} - \Sigma c_s \frac{\partial G'}{\partial c_s}\right], \quad (25)$$

and since, $\Sigma m_{01} c_{01}$ being unity, $\Sigma m_{01} dc_{01}$ is always zero, we have

$$G' - G = \int_{C=0} \Sigma \frac{\partial G'}{\partial c_{01}} dc_{01}. \quad . \quad . \quad . \quad . \quad (26)$$

To say that any such modification of the division into "linear" and "general" terms is due to, and represents, a departure on the part of the solvent molecules from "perfect" behaviour would be to make a qualitative statement of no

practical significance. (24) is a quantitative statement, and the form to which it is reduced in (26) has practical meaning and value, as will be shown by means of a simple illustration in a short appendix to this section. From (26) it is clear that, owing to the relative smallness of the changes in c_{01}, c_{02}, ..., produced by the presence of solutes, the difference between the "general" terms in the original "molecular" expression for ψ and those in the "experimental" expression we have obtained will generally be very small. But it may have to be taken into account in making use of dynamical theory (at high concentrations).

Of course, until and except when the "general" terms can be given more definite form, we cannot say anything about the way in which they will depend on the constitution of the solvent and its variation. For the present we have to suppose that $c_{01}c_{02}$ ..... are eliminated in terms of T, p, c_1, c_2, ..., from G, which takes some form

$$G = R\Sigma\phi_x' f_x(c_1 c_2 \ldots\ldots_s). \quad . \quad . \quad . \quad . \quad (27)$$

Theoretically, and in the general case, the application of dynamical theory will precede this elimination of $c_{01}c_{02}$....., will deal in fact with the original general terms M_0G', so that, in greater or less degree, knowledge of the constitution of the solvent and its variation will be necessary before such theory can join issue with practice, but in some cases, as the appendix will illustrate, this may not be necessary, even though the constitution of the solvent does affect the "general" terms.

It should be noticed that the value and convenience of our "experimental" expression for ψ is by no means entirely dependent on a deficiency of knowledge of the constitution of the solvent, though the latter makes it practically indispensable.

In the "linear" terms the effect of the constitution of the solvent is concentrated in the quantity $\overline{m}$, the mean molecular weight of the solvent.

In the pure solvent this will have a limiting value $\overline{m}_0$ dependent only on temperature and pressure, and we can therefore regard the quantity $(-R \log \overline{m}_0)$ as included in the quantity ϕ_s, when our "experimental" expression for ψ will finally take the form

$$\psi = M_0\left[\phi_M + R\int_{C=0}^{C}\frac{1}{\overline{m}}\,d\log(1+\overline{m}C)\right]$$

$$+\Sigma n_s\left[\phi_s - R\left\{\log c_s - \log\left(\frac{\overline{m}_0}{\overline{m}} + \overline{m}_0 C\right\}\right] + M_0 G \quad . \quad . \quad (28)$$

when $\overline{m}$ is constant (at constant temperature and pressure) and therefore equal to $\overline{m}_0$ we get

$$\psi = M_0\left[\phi_M + \frac{R}{\overline{m}_0}\log(1+\overline{m}_0 C)\right]$$

$$+\Sigma n_s\left[\phi_s - R\left\{\log c_s - \log(1+\overline{m}_0 C)\right\}\right] + M_0 G, \quad . \quad . \quad (20)$$

which is equation (52) of the first of these papers, from which the "second approximation" equations were obtained.

It is, of course, not possible to say how, in the most general case, $\overline{m}$ will depend on the concentrations of the various solutes, but an interesting case, of probably very wide application, may be treated and will at the same time serve as an illustration.

This is the case where $\overline{m}$ can be written, with sufficient approximation, as a series of ascending integral powers of C, the total solute concentration.

This can be shown to be the case, for instance, when the various solvent-molecular species behave as perfect solutes (in the true sense,—not in the sense of the Raoult-van't Hoff laws).

Some simple cases have been investigated, but the detail need not be given here. It will suffice to say that in the simplest case, for example where the solvent consists of two molecular species, m_0 and $(2\,m_0)$, the one the doublet of the other, we find that $\overline{m}$ can be expressed as

$$\overline{m} = \overline{m}_0[1+\theta(\overline{m}_0 C)+\eta(\overline{m}_0 C)^2],$$

where the values of θ and η depend, of course, on the proportions in which the two species are present in the pure solvent, but *in any case cannot exceed* $\frac{1}{8}$ and $\frac{1}{10}$ respectively (these maxima not being simultaneous).

We may carry this expansion of $\overline{m}$, which is formally convenient for our purpose, to one further term of which only the order of magnitude will matter,

$$\frac{\overline{m}}{\overline{m}_0} = 1+\theta(\overline{m}_0 C)+\eta(\overline{m}_0 C)^2+\xi(\overline{m}_0 C)^3, \quad . \quad . \quad (30)$$

the last term being, as we shall find, altogether negligible if ξ is no greater than about $\frac{1}{10}$.

Approximating on the assumption that θ and η are of the

order of magnitude $\frac{1}{10}$, we obtain :

$$\log\left(\frac{\overline{m}_0}{\overline{m}}+\overline{m}_0\mathrm{C}\right)=a_1\mathrm{C}-\tfrac{1}{2}a_2\mathrm{C}^2+\tfrac{1}{3}a_3\mathrm{C}^3, \quad . \quad . \quad (31)$$

$$\int_{\mathrm{C}=0}^{\frac{1}{\overline{m}}} d\log(1+\overline{m}\mathrm{C})=\mathrm{C}(1-\tfrac{1}{2}a_1\mathrm{C}+\tfrac{1}{3}a_2\mathrm{C}^2-\tfrac{1}{4}a_3\mathrm{C}^3), \quad (32)$$

where

$$\left.\begin{aligned} a_1&=\overline{m}_0(1-\theta),\\ a_2&=\overline{m}_0{}^2[1-2(\theta-\eta)],\\ a_3&=\overline{m}_0{}^3[1-3(\theta-\eta+\xi)]. \end{aligned}\right\} \quad . \quad . \quad . \quad (33)$$

it being clear that, as stated above, ξ can be neglected altogether if no greater than about $\frac{1}{10}$.

We get then the equations

$$\text{IV.}\quad \left\{\begin{aligned} \frac{\partial\psi}{\partial \mathrm{M}_0}&=\phi_{\mathrm{M}}+\mathrm{RC}[1-\tfrac{1}{2}a_1\mathrm{C}+\tfrac{1}{3}a_2\mathrm{C}^2-\tfrac{1}{4}a_3\mathrm{C}^3]+\mathrm{G}_{\mathrm{M}}\\ \frac{\partial\psi}{\partial n_s}&=\phi_s-\mathrm{R}[\log c_s-a_1\mathrm{C}+\tfrac{1}{2}a_2\mathrm{C}^2-\tfrac{1}{3}a_3\mathrm{C}^3]+\mathrm{G}_s \end{aligned}\right\} (34)$$

and thence the successive stages :—

$$\text{III.}\quad \left\{\begin{aligned} \frac{\partial\psi}{\partial \mathrm{M}_0}&=\phi_{\mathrm{M}}+\mathrm{RC}[1-\tfrac{1}{2}\overline{m}_0(1-\theta)\mathrm{C}+\tfrac{1}{3}\overline{m}_0{}^2\mathrm{C}^2]+\mathrm{G}_{\mathrm{M}}\\ \frac{\partial\psi}{\partial n_s}&=\phi_s-\mathrm{R}[\log c_s-\overline{m}_0(1-\theta)\mathrm{C}+\tfrac{1}{2}m_0{}^2\mathrm{C}^2]+\mathrm{G}_s \end{aligned}\right\} (35)$$

and

$$\text{II.}\quad \left\{\begin{aligned} \frac{\partial\psi}{\partial \mathrm{M}_0}&=\phi_{\mathrm{M}}+\mathrm{RC}[1-\tfrac{1}{2}\overline{m}_0\mathrm{C}]+\mathrm{G}_{\mathrm{M}}\\ \frac{\partial\psi}{\partial n_s}&=\phi_s-\mathrm{R}[\log c_s-\overline{m}_0\mathrm{C}]+\mathrm{G}_s \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (36)$$

and

$$\text{I.}\quad \left\{\begin{aligned} \frac{\partial\psi}{\partial \mathrm{M}_0}&=\phi_{\mathrm{M}}+\mathrm{RC}+\mathrm{G}_{\mathrm{M}}\\ \frac{\partial\psi}{\partial n_s}&=\phi_s-\mathrm{R}\log c_s+\mathrm{G}_s \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

I. and II. being the first and second approximations obtained by the preliminary treatment in the previous paper.

Taking $\frac{1}{5}$ per cent. as the "probable experimental error,"

we get roughly the following upper limits (of total concentration C) for the applicability of the four successive stages of approximation :

$$\left.\begin{array}{lll} \text{Approximation I.....} & \dfrac{\text{M.}}{10} & \\ \text{,,} \quad \text{II.....} & 2\ \text{M.} & \\ \text{,,} \quad \text{III.....} & 5\ \text{M.} & \\ \text{,,} \quad \text{IV.....} & 8\ \text{M.} & \end{array}\right\} \quad . \quad . \quad (38)$$

—that is, considering aqueous solution, and assuming $\overline{m}_0$ to be about 40. The limits would be considerably different, particularly in the case of III. and IV., if $\overline{m}_0$ were given a very different value, as can readily be seen.

The practical criterion of a perfect solute in a complex solvent is now that its behaviour should be expressed by that one of the above succesive approximations appropriate to the total-solute concentration of the solution, *with the "general" terms omitted.*

If the assumption of perfect behaviour in the case of a particular solute be made, an experimental determination of the quantities $\overline{m}_0$, θ, η, etc., can be made and concordance in several such determinations made upon different solutes would tend to justify the assumption that perfect behaviour persisted up to the concentration at which concordance was found.

According to the concentration reached (with concordance) some of the quantities $\overline{m}_0$, θ, η, etc. would then be known with some approximation (closest in $\overline{m}_0$, next in θ, and so on).

Then, on the assumption of "perfect" behaviour on the part of the solvent-molecular species—that is, a sufficient approximation thereto,—these quantities would suffice to discover something about the constitution of the solvent.

Thus two solvent-molecular species would be completely determined by a knowledge of $\overline{m}_0$ alone (that is, the proportions of the two kinds present would be determined),—and in this case the question of "perfect" behaviour of the two species would not enter. $\overline{m}_0$ and θ would suffice to determine completely three species, or would provide a check if only two species were present, and so on.

The problem, however, as a practical problem is complicated by considerations which are the subject of a succeeding paper, viz. the question of solvation and partial solvation of solutes.

Finally, it is proposed to consider one point with regard

to the expansions of U and V, the total energy and volume of the solution. (Fuller consideration is postponed, as this paper is already rather long.)

In "perfect" solution we have

$$\left.\begin{aligned} \mathrm{U} &= \Sigma n_{01}u_{01} + \Sigma n_s u_s \\ \mathrm{V} &= \Sigma n_{01}v_{01} + \Sigma n_s v_s \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (39)$$

and

$$\mathrm{Q} = \Sigma n_{01}q_{01} + \Sigma n_s q_s, \quad . \quad . \quad . \quad . \quad . \quad (40)$$

where, according to the usage of the previous paper, Q is $(\mathrm{U}+p\mathrm{V})$, q_s is (u_s+pv_s), etc., and we can write this

$$\mathrm{Q} = \mathrm{M}_0\Sigma c_{01}q_{01} + \Sigma n_s q_s. \quad . \quad . \quad . \quad . \quad (41)$$

$\Sigma c_{01}q_{01}$, however, depends (through $c_{01}, c_{02}; \ldots.$) upon the concentrations of solutes present, but has, in the pure solvent, a limiting value q_M depending only on T and p, and

$$\Sigma c_{01}q_{01} = q_\mathrm{M} + \Sigma q_{01}\int_{\mathrm{C}=0}^{\mathrm{C}} dc_{01} \quad . \quad . \quad . \quad . \quad (42)$$

$$= q_\mathrm{M} + \Sigma q_{01}\Delta c_{01}. \quad . \quad . \quad . \quad . \quad . \quad (43)$$

And so

$$\mathrm{Q} = \mathrm{M}_0 q_\mathrm{M} + \Sigma n_s q_s + \mathrm{M}_0\Sigma q_{01}\Delta c_{01}. \quad . \quad . \quad . \quad (44)$$

Dilute now such a solution "infinitely" by adding a large mass M_0' of pure solvent at the same T and p, for which

$$\mathrm{Q}' = \mathrm{M}_0' q_\mathrm{M}.$$

The united heat will be

$$(\mathrm{Q}+\mathrm{Q}') = (\mathrm{M}_0+\mathrm{M}_0')q_\mathrm{M} + \Sigma n_s q_s + \mathrm{M}_0\Sigma q_{01}\Delta c_{01}.$$

But since the solution is now "infinitely" dilute, we shall have on bringing it to the original temperature and pressure

$$\mathrm{Q}'' = (\mathrm{M}_0+\mathrm{M}_0')q_\mathrm{M} + \Sigma n_s q = (\mathrm{Q}+\mathrm{Q}') - \mathrm{M}_0\Sigma q_{01}\Delta c_{01}.$$

In other words, $\mathrm{M}_0\Sigma q_{01}\Delta c_{01}$ was the heat developed (evolved) on diluting the Σn_s molecules of solute.

That is, there is a *heat of dilution* of

$$\frac{1}{\mathrm{C}}\Sigma q_{01}\Delta c_{01} \quad . \quad . \quad . \quad . \quad . \quad . \quad (45)$$

per gram-molecule of solute, in "*perfect*" *solution* in a "complex" solvent. The explanation of this apparent

anomaly is that in a "complex" solvent the process of dilution is not quite so simple as in a "simple" solvent, being accompanied by a change in the constitution of the solvent—a reversion, in fact, to the constitution of the pure solvent. It is, of course, plain that $M_0\Sigma q_{01}\Delta c_{01}$, or $\Sigma q_{01}\Delta n_{01}$, is the ("chemical") heat evolution accompanying this reversion (at constant temperature and pressure).

With (44) analogously we have

$$U = M_0 u_M + \Sigma n_s u_s + M_0 \Sigma u_{01}\Delta c_{01}, \quad . \quad . \quad . \quad (46)$$

$$V = M_0 v_M + \Sigma n_s v_s + M_0 \Sigma v_{01}\Delta c_{01}, \quad . \quad . \quad . \quad (47)$$

and it may easily be verified from (11) that

$$\left.\begin{aligned} u_M &= T^2 \frac{\partial \phi_M}{\partial T}, \\ v_M &= -T \frac{\partial \phi_M}{\partial p}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (48)$$

Appendix to Section III.

A simple example, which is essentially merely illustrative, but may possibly be something more, will serve to make clear the practical significance of equation (26) concerning the slight modification of the "general" terms.

In section V. of the first of these papers was obtained for the general term in the case of a dilute solution of a binary strong electrolyte

$$M_0 G' = R M_0 \phi' c_s^{3/2}, \quad . \quad . \quad . \quad . \quad . \quad (49)$$

where ϕ' depended, in some way, upon a certain "effective" dielectric constant (D), which, at sufficient dilution, would be that of the pure solvent.

Let us suppose that D depends on c_s in such a way that ϕ' is a linear function of c_s

$$\phi' = \phi_0'(1 + \alpha c_s), \quad . \quad . \quad . \quad . \quad . \quad (50)$$

and also, in the first instance, that this effect of c_s on D is entirely independent of its effect on the constitution of the solvent, that is, that the slight change in the latter produced by the addition of the electrolyte would *alone* [if produced, for instance, by some different solute, c_s, the concentration

of the electrolyte being made quite small] have but negligible effect on D, then G′ is independent of the constitution of the solvent (practically), *i. e.*

$$\int_{C=0} \Sigma \frac{\partial G'}{\partial c_{01}} dc_{01} = 0 ; \quad . \quad . \quad . \quad . \quad . \quad (51)$$

so that,

$$\left.\begin{aligned} G &= G' &&= R\phi_0' c_s^{3/2}(1+\alpha c_s) \\ G_s &= G_s' &&= \tfrac{3}{2} R\phi_0' c_s^{1/2}(1+\tfrac{5}{3}\alpha c_s) \\ G_M &= \Sigma c_{01} G_{01}' &&= -\tfrac{1}{2} R\phi_0' c_s^{3/2}(1+3\alpha c_s) \end{aligned}\right\} . \quad . \quad (52)$$

If, on the other hand, we suppose that the effect of c_s on D is entirely dependent on the change produced in the constitution of the solvent, and would be fully obtained if the latter could be brought about in some other way, while c_s was made very small, then we should have

$$\left.\begin{aligned} \Sigma \frac{\partial G'}{\partial c_{01}} \cdot \frac{\partial c_{01}}{\partial c_s} &= R c_s^{3/2} \frac{\partial \phi'}{\partial c_s} \\ &= R c_s^{3/2} \phi' \alpha \end{aligned}\right\} \quad . \quad . \quad . \quad (53)$$

and

$$\int_{C=0} \Sigma \frac{\partial G'}{\partial c_{01}} dc_{01} = R\phi_0' \alpha \int_{c_s=0} c_s^{3/2} d \, . \, c_s, \quad . \quad . \quad . \quad (54)$$

so that

$$G' - G = \tfrac{2}{5} R\phi_0' \alpha c_s^{5/2}, \quad . \quad . \quad . \quad . \quad (55)$$

i. e. in this second case,

$$\left.\begin{aligned} G &= G' &&- \int_{C=0} \Sigma \frac{\partial G'}{\partial c_{01}} dc_{01} = R\phi_0' c_s^{3/2}(1+\tfrac{3}{5}\alpha c_s) \\ G_s &= G_s' &&\qquad\qquad = \tfrac{3}{2} R\phi_0' c_s^{1/2}(1+\alpha c_s) \\ G_M &= \Sigma c_{01} G_{01}' &&- \int_{C=0} \Sigma \frac{\partial G'}{\partial c_{01}} dc_{01} = -\tfrac{1}{2} R\phi_0' c_s^{3/2}(1+\tfrac{9}{5}\alpha c_s) \end{aligned}\right\} . \quad (56)$$

In both cases, of course,

$$G_s = G_s',$$

but they are not the same in the two cases because in the second case ϕ' is a constant when $n_{01}, n_{02} \ldots .$ are held constant, but not in the first case. (For the same reason $\Sigma c_{01} G_{01}'$ is also different in the two cases.)

In the second case we note that although the constitution

of the solvent enters into the "general" terms it does so only through the quantity D, and if this is the dielectric constant of the solution in bulk it can be measured and so determined as a function of c_s without considering the corresponding variation in the constitution of the solvent, or the way in which the latter exerts its effect on D.

If, as is probable at the less extreme dilutions, D is not the experimental or "bulk" dielectric constant, but a certain statistical-average quantity of a peculiar kind, then its variation is, at least partially, not due to a variation in the solvent, but directly to variation of c_s as in the first case above.

Clearly from the preceding it would theoretically be possible in such a case to determine from comparison of theory with experiment whether the effect of c_s on D was direct or indirect or in what proportion both, but it might not be practicable owing to the smallness of the effects to be measured.

Note on the "Gibbs Fundamental Relation."

Consider any property π of a homogeneous substance or phase, which is determined in magnitude by the composition of the phase and the quantity of substance considered. In view of the homogeneity it must then be *proportional* to the quantity of substance when the composition is fixed.

Such properties are (at constant temperature and pressure) U and V, the total-energy and volume, Q or $(U+pV)$, which might be called the "reversible heat content," and any thermodynamic potential such as entropy, free energy, Gibbs' "chemical potential," or Planck's ψ, which may all be expressed (at constant temperature and pressure) as functions of the quantities $M_1 M_2 \ldots$ of the constituents which suffice, under the conditions considered, to produce the phase.

We can show, as Planck does in the case of ψ, that for any such property π,

$$\pi = \Sigma M_1 \frac{\partial \pi}{\partial M_1}, \quad \ldots \ldots \quad (57)$$

for if ϵ be some infinitesimal fraction and we remove (ϵM_1) of the first constituent, clearly π is diminished by $\epsilon M_1 \frac{\partial \pi}{\partial M_1}$.

Removing simultaneously the same fraction of the total quantity of each constituent we diminish π, in all, by $\Sigma \epsilon M_1 \frac{\partial \pi}{\partial M_1}$, or $\epsilon \Sigma M_1 \frac{\partial \pi}{\partial M_1}$. But, in so doing, we have

simply removed the fraction ϵ of the whole phase, without altering its composition, so that π must have diminished by $\epsilon\pi$, that is

$$\epsilon\pi = \epsilon\Sigma M_1 \frac{\partial \pi}{\partial M_1},$$

whence (57) follows.

When π is either ψ, or Gibbs' "chemical potential," a whole system of phases *in equilibrium* can be considered together, since then $\frac{\partial \pi}{\partial M_1}$, etc., are the same in every phase.

Without actually quoting Euler's theorem, Planck remarks, in regard to ψ, that this relation means that ψ, as a function of $M_1 M_2 \ldots$, is homogeneous and of the first degree, though not, of course, in general linear, and the same remark applies to our typical property π.

From (57) we can at once get a more practically useful relation by differentiating both sides fully:—

$$\Sigma \frac{\partial \pi}{\partial M_1} dM_1 = \Sigma \frac{\partial \pi}{\partial M_1} dM_1 + \Sigma M_1 d\left(\frac{\partial \pi}{\partial M_1}\right),$$

that is,

$$\Sigma M_1 d\left(\frac{\partial \pi}{\partial M_1}\right) = 0. \quad . \quad . \quad . \quad . \quad . \quad (58)$$

Now equation (97) of Gibbs' classical paper reduces at constant temperature and pressure to

$$\Sigma m_1 d\mu_1 = 0$$

and is then simply (58) applied to Gibbs' "chemical potential."

(58) is therefore referred to in these papers as the "Gibbs fundamental relation," but its general applicability to any property of the type of π (for a single homogeneous phase) is to be borne in mind.

It is to be observed that while the constituents whose masses are $M_1 M_2 \ldots$ must be sufficient to produce the phase under the conditions considered, they need not be all necessary—they need not be the "general-thermodynamic" components. And since also (obviously from the form of (58)) $M_1 M_2 \ldots$ need not be expressed in the same units we see that equally valid is the form

$$\Sigma n_1 d\left(\frac{\partial \pi}{\partial n_1}\right) = 0, \quad . \quad . \quad . \quad . \quad . \quad (59)$$

the "molecular-thermodynamic" form, in which $n_1 n_2 \ldots$ are the numbers present of the various molecular species.

It is an important point in the treatment of the two problems, "complex solvents" and "solvation," presented in this and a succeeding paper, that in the "practical" or "experimental" expression for ψ finally obtained the "linear terms" by themselves satisfy the Gibbs fundamental relation, for in perfect solution these terms alone remain. And this is preserved in the successive approximations. The relation also serves as a useful check upon the correctness of the detail.

Balliol College,
March 1922.

XVII. *The Calculation of Centroids.* *By* J. G. Gray, *D.Sc.*, *Cargill Professor of Applied Physics in the University of Glasgow* *.

THE position of the centroid of a plane arc or area is usually determined by the application of one or other of the two theorems of Pappus. The methods described and illustrated below seem, however, to be novel; they are useful in a great number of cases, including many to which the theorems of Pappus do not apply.

Fig. 1.

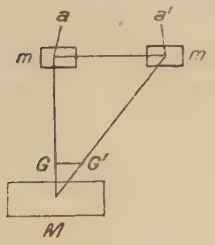
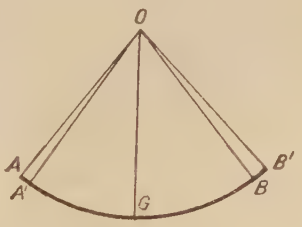


Fig. 2.



Consider a system made up of two masses M and m (fig. 1). Let the centroids of m and of the system M and m be at a and G respectively. Now suppose the mass m moved so that its centroid is brought to a'. G moves to G', where GG' is parallel to aa'; and we have $(M+m)\,\overline{GG'} = m\,\overline{aa'}$.

As a first example, consider the case of a circular arc AB mass per unit of length m (say). Let O be the centre of the circle of which the arc forms part. Now suppose the arc

* Communicated by the Author.

rotated about O in its own plane through a small angle θ, so that A is brought to A′ and B to B′. In effect the small portion AA′ of the arc is transferred from one end of the arc to the other. The mass of this element is $mr\theta$, and it has been translated (virtually) through the distance $2r \sin \alpha/2$, where α is the angle AOB. If G is the centroid of the arc, we have obviously

$$mr\theta \, 2r \sin \alpha/2 = mr\alpha \, \overline{\mathrm{OG}} \, \theta,$$

or

$$\mathrm{OG} = \frac{2r \sin \alpha/2}{\alpha}.$$

As a second example we take the case of a sector of a circle OADC (fig. 3). Let the sector be turned in its own plane through a small angle θ about an axis through O, so that A comes to A′, B to B′. The effect of the rotation has been (virtually) to transfer the triangle OAA′ to OCC′.

Fig. 3.

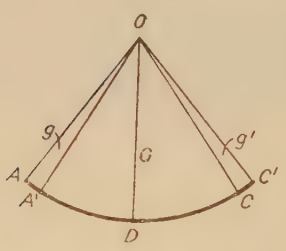
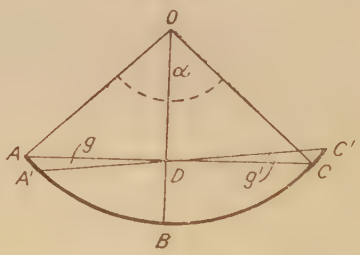


Fig. 4.



The centroid G of the sector has moved parallel to gg' through a distance $\overline{\mathrm{OG}} \, \theta$. The mass of the sector is $\frac{1}{2}r^2\alpha\sigma$, and that of the triangle OAA′ is $\frac{1}{2}r^2\theta\sigma$, where σ is the mass of the sector per unit area. Since $gg' = \frac{4}{3} r \sin \frac{\alpha}{2}$, we have

$$\frac{1}{2}r^2\theta\sigma \frac{4}{3} r \sin \frac{\alpha}{2} = \frac{1}{2} r^2 \alpha\sigma \, \overline{\mathrm{OG}} \, \theta,$$

or

$$\overline{\mathrm{OG}} = \frac{4}{3} \frac{r \sin \alpha/2}{\alpha}.$$

Again, let it be required to find the position of the centroid of a segment of a circle ABC (fig. 4). The segment is turned in its own plane, about O, through a small angle θ. A is thus brought to A′ and C to C′. If the mass per unit of area of the segment is σ, the mass of the triangle DAA′

is $\sigma\theta \frac{1}{2} r \sin \frac{1}{2}\alpha \, r \sin \frac{1}{2}\alpha$, or $\frac{1}{2}\sigma\theta r^2 \sin^2 \frac{1}{2}\alpha$. The area of the segment is $\frac{1}{2}r^2\alpha - r^2 \sin \frac{1}{2}\alpha \cos \frac{1}{2}\alpha$, and its mass is

$$\sigma r^2(\tfrac{1}{2}\alpha - \sin \tfrac{1}{2}\alpha \cos \tfrac{1}{2}\alpha).$$

Since gg' is $\frac{4}{3} r \sin \frac{1}{2}\alpha$, we have

$$\sigma r^2(\tfrac{1}{2}\alpha - \sin \tfrac{1}{2}\alpha \cos \tfrac{1}{2}\alpha) x\theta = \tfrac{2}{3}\sigma\theta r^3 \sin^3 \tfrac{1}{2}\alpha,$$

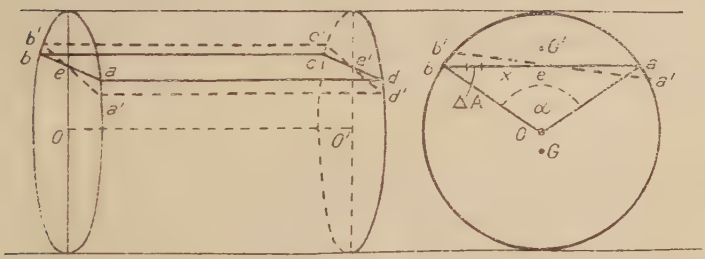
or

$$x = \frac{2}{3} \frac{r \sin^3 \frac{1}{2}\alpha}{\frac{1}{2}\alpha - \sin \frac{1}{2}\alpha \cos \frac{1}{2}\alpha}.$$

where x is the distance of the centroid of the segment from O.

Consider next the solids obtained by dividing a right circular cylinder into two parts by means of the plane *abcd* (fig. 5). Let it be required to find the position of the

Fig. 5.



centroid of the lower solid. We suppose the solid rotated through a small angle θ about the axis OO′ (the axis of figure of the complete cylinder); a is thus brought to a', b to b', c to c', and d to d'. In effect the wedge $ebb'e'cc'$ has been removed from the solid and replaced in the position $eaa'edd'$. If ΔA denotes an element of area in *abcd* at a distance x from ee', the volume swept out by this element due to the turning of the solid is $\Delta A x\theta$. The mass of this element of volume is $\rho \Delta A x\theta$, and since the element is moved (virtually) through a distance $2x$, we have, if V is the volume of the solid,

$$V\rho \, \overline{OG}\, \theta = 2\rho\theta \Sigma \Delta A x^2,$$

where the summation is made over the complete area *abcd*. Hence

$$V \times \overline{OG} = AK^2,$$

where A is the area *abcd*, and K is its radius of gyration

about ee'. If α is the angle $bO\alpha$, and l the length of the solid, we have for the sectional area

$$\tfrac{1}{2}r^2(2\pi-\alpha)+r^2\sin\tfrac{1}{2}\alpha\cos\tfrac{1}{2}\alpha.$$

And since $A = 2lr\sin\frac{1}{2}\alpha$, and $K^2 = \frac{1}{3}r^2\sin^2\frac{1}{2}\alpha$, we have

$$OG = \frac{\frac{2}{3}r\sin^3\frac{1}{2}\alpha}{\pi-\frac{1}{2}\alpha+\sin\frac{1}{2}\alpha\cos\frac{1}{2}\alpha}.$$

Similarly, if G′ is the centroid of the upper solid, we have

$$OG' = \frac{\frac{2}{3}r\sin^3\frac{1}{2}\alpha}{\frac{1}{2}\alpha-\sin\frac{1}{2}\alpha\cos\frac{1}{2}\alpha}.$$

The positions of the centroids of the solids obtained by dividing a right circular cylinder into two portions by means of planes $O\alpha$, Ob (fig. 6) are easily determined. If G is the centroid of the larger solid, we obtain at once, by supposing

Fig. 6.

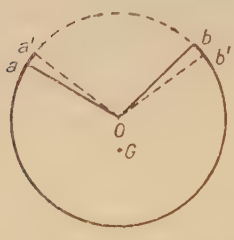
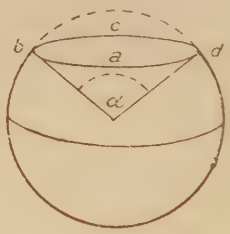


Fig. 7.



the solid turned through a small angle θ about the axis of figure of the complete cylinder, so that α arrives at α', and b at b',

$$V\times\overline{OG} = 2AK^2\sin\tfrac{1}{2}\alpha,$$

where A is the area represented by $O\alpha$, and K is its radius of gyration about the axis of figure. We have, if l is the length of the cylinder,

$$OG = \frac{2lr\frac{1}{3}r^2\sin\frac{1}{2}\alpha}{\frac{1}{2}r^2(2\pi-\alpha)l}$$

$$= \frac{4}{3}\frac{r\sin\frac{1}{2}\alpha}{2\pi-\alpha}.$$

If the cylinder is divided into two equal parts, we obtain

the position of the centroid of each by putting $\alpha = \pi$. Thus

$$OG = \frac{4}{3}\frac{r}{\pi}.$$

For the portion of a sphere shown by the firm lines of figure 7, we have, if OG is the distance of the centroid from the centre of the complete sphere,

$$OG = \frac{AK^2}{V},$$

where A is the area *abcd*, and K is its radius of gyration about its diameter. Thus

$$OG = \frac{\pi r^2 \sin^2 \frac{1}{2}\alpha \frac{1}{4}r^2 \sin^2 \frac{1}{2}\alpha}{\frac{1}{3}r^2(1+\cos\frac{1}{2}\alpha) \times 2\pi r}$$

$$= \frac{3}{8}\frac{r}{1+\cos\frac{1}{2}\alpha}.$$

For the portion of a sphere enclosed by the firm lines of figure 8 we have

$$x = \frac{\pi r^2 \frac{1}{4}r^2 \sin\frac{1}{2}\alpha}{\frac{4}{3}\pi r^3 \frac{2\pi-\alpha}{2\pi}}$$

$$= \frac{3}{8}\frac{\pi r \sin\frac{1}{2}\alpha}{2\pi-\alpha},$$

where x is the distance of the centroid from OO′. For a hemisphere $\alpha = \pi$, and

$$x = \tfrac{3}{8}r.$$

Finally, for the portion of an anchor ring shown in figure 9 we have, if r is the radius of the ring and a the radius of the section,

$$OG = \frac{\pi a^2(\frac{1}{4}a^2 + r^2)\, 2\sin\frac{1}{2}\alpha}{\pi a^2 \times 2\pi r \times \frac{2\pi-\alpha}{2\pi}}$$

$$= \frac{2(r^2+\frac{1}{4}a^2)\sin\frac{1}{2}\alpha}{r(2\pi-\alpha)},$$

which reduces to $2r/\pi$ when $\alpha = \pi$ and a is very small in comparison with r.

If a body is floating partly immersed in a liquid, the

distance of the centre of buoyancy B (the centroid of the displaced liquid) from the metacentre M is given by $BM = AK^2/V$, where V is the volume of the displaced liquid and AK^2 is the moment of inertia of the water-line area about the intersection of the wedges of emersion and immersion. The equilibrium of the floating body is stable if

Fig. 8.

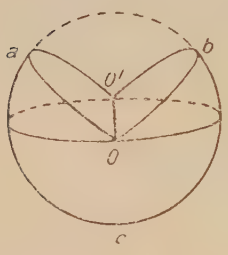
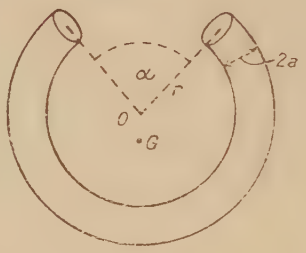


Fig. 9.



M is above G, the centroid of the floating body, that is if $BM > BG$; it is unstable if $BM < BG$; and it is neutral if $BM = BG$, that is when M coincides with G. In figure 5 we may suppose the complete cylinder floating in water so that *abcd* is the water-line area. The cylinder is obviously in neutral equilibrium so far as turning about the axis OO′ is concerned. Thus M lies in the line OO′, and we have

$$OG = \frac{AK^2}{V}.$$

University of Glasgow,
May 1, 1922.

XVIII. *Preliminary Investigations on Silberstein's Quantum Theory of Photographic Exposure.* *By* A. P. H. Trivelli *and* L. Righter *.

Introductory.

THIS paper is the first of a number of investigations now being conducted in this laboratory to test experimentally the light-quantum theory of photographic exposure recently proposed by Dr. Silberstein before the Toronto meeting of the American Physical Society.

* Communication No. 141 from the Research Laboratory of the Eastman Kodak Company.

Originally, these experiments were started independently of that theory, our intention being primarily to study the effect of the clumping, or clustering together in groups, of silver halide grains in a photographic emulsion *. According to Slade and Higson †, "It seems reasonable to assume that each grain acts quite independently and that one grain which has become developable is unable to make a grain, situated in close proximity to it, developable unless the latter grain is developable in itself." From the table (p. 256) it is readily seen that this statement is not true, but the grains when clumped together *act as one grain* for development to the limit.

Since each clump acts as one grain a very much broader range of grain sizes, or their equivalent, is obtained extending from the smallest single grains up to the largest clumps (containing 30 or more grains) in a given emulsion.

It was thought that these results afforded a rigorous test for Silberstein's theory, and it seemed therefore worth while to compare them with the implications of that theory.

Silberstein's fundamental formula is essentially, *i. e.* apart from chromatic complications, and disregarding the lateral dimensions of the light quanta,

$$\frac{k}{N} = 1 - e^{-na},$$

where N is the original number, per unit of the plate, of the class of grains of size (area) a, n the number of incident light-quanta, again per unit area of plate, and k the number of grains affected ‡. If the finite cross section σ of the light-quantum is taken into account, a is to be replaced by

$$a' = a\left[1 - \sqrt{\frac{\sigma}{a}}\right]^2.$$

At first, the rapid increase of k with the size a as required by that formula seemed (qualitatively speaking) not only attained but even exceeded by the experimental results. This seemed to us to indicate that the sizes (areas) of the clumps of grains were under-estimated by us. In fact, for

* Extensive experiments are being conducted in this laboratory to study the clumping of grains in different concentrations of the same emulsion. These investigations will be published later.

† Slade and Higson, "Photochemical Investigations of the Photographic Plate." Proc. Roy. Soc. xcviii. p. 154 (1920).

‡ Silberstein, L., "Quantum Theory of Photographic Exposure," *infrà*, p. 257.

an estimate of the areas of all clumps one and the same average grain size (area) had been assumed throughout. Upon recalculating the results, however, and assigning the correct average grain size (area) to the single grains and to the different clumps, the very interesting fact was observed that the average grain size (area) increases from the single grains to the clumps of two, three, etc. The corrected results conform, even better than expected, to the above formula, with a finite σ.

Experimental.

These experiments, although of an extremely tedious and trying nature, were performed with the utmost care and, to the best of our knowledge, all sources of error were either eliminated or reduced to a minimum. Only a brief description of the experimental procedure will be given at this time, as a more extensive paper containing further experimental results is to be published in the near future with a detailed account of our methods of photomicrography, and in which all errors will be discussed fully.

A simple silver bromide emulsion was used for these experiments having a speed of 112 and gamma 0·8 for six minutes development in an ordinary pyro-soda developer at 17° C. The average size of the grain is about 0·9 μ diameter.

The method of preparation of strips for sensitometric exposure is, briefly, as follows:—One 5 in. × 7 in. plate of the original emulsion is soaked in distilled water for one half-hour at 0° C. to 8° C. (All work for sensitometric exposure was done in a dark room by the aid of a dull red safelight, Wratten Series 2). The water is removed and a warm solution of gelatine, alcohol, and water is added and the whole solution heated in an oven for 20 minutes, while care is being taken not to heat over 40° C., because above this temperature much fogging takes place. With several such applications of the aforesaid solution the emulsion is entirely removed from the plate and the resulting solution made up to such a volume that it will give one layer of grains upon coating and drying. Some of the slides are used at once to get the clump frequency data. Those for exposure are backed with an opaque substance to prevent reflexions, then exposed in a sensitometer, developed to gamma infinity with a pyro-soda developer at 17° C., washed, and the developed silver removed with a dilute solution of chromic acid and

sulphuric acid. The strips thus obtained contain the undeveloped grains, and by taking the difference the number of developed grains is calculated.

The data given below cover only the first or highest density step of a Hurter and Driffield sensitometric strip *. 20 fields on each of 3 strips are employed to determine the developed grains. To determine the number of grains and clumps in the original one grain layer plate before sensitometric exposure and development, 10 fields on each of 4 strips were used. By taking this large number of fields on several strips we obtained a much better average. The results in both the above cases are reduced to a number of grains or clumps per square centimetre of one grain layer plate. Then as the dilution is known, one may, with certain restrictions, refer back to the original plate.

All photomicrographs were made at a magnification of 2500 diameters and these negatives enlarged 4 times in printing. On the prints the grains and clumps are measured and counted, and then classified in class sizes (areas). The class sizes (areas) are 0 to $0{\cdot}2\,\mu^2$, $0{\cdot}2$ to $0{\cdot}4\,\mu^2$, $0{\cdot}4$ to $0{\cdot}6\,\mu^2$, etc. The light source is a point source from a Pointolite lamp which is screened with a Wratten (H) blue filter to restrict the wave-length range and therefore increase the resolving power of the microscope. A cell containing copper sulphate solution absorbs heat rays and a cell containing a solution of quinine bisulphate excludes the ultraviolet light. The optical system is built up as follows: Cedar oil immersion condenser and objective, aplanatic condenser of numerical aperture 1·4, and Bausch and Lomb objective 1·9 mm. numerical aperture 1·3 in combination with a No. 6 compensating ocular.

In the following is a table of our results. Column 1 contains the number of grains in each clump, column 2 the average area of grains in corresponding clumps, column 3 the number of grains times 10^{-3} per square centimetre of original one-grain layer plate, column 4 the number of grains times 10^{-3} per square centimetre of developed one-grain layer plate, and column 5 gives the proportionate number $\frac{k}{N}$ of clumps affected.

* The five remaining steps each corresponding to one-half the exposure of the preceding are now being counted and mapped out, and we hope to be able to publish the results obtained with them in the course of one or two months.

Number of Grains in Clump.	a in μ^2.	$N \cdot 10^{-3}$ per sq. cm. 1 grain layer plate (original).	$k \cdot 10^{-3}$ per sq. cm. 1 grain layer plate (developed).	$\frac{k}{N}$
1	0·75[4]	6577·0	1086·0	0·165
2	1·92[5]	1322·0	594·0	0·449
3	3·03	664·0	508·6	0·766
4	4·88[1]	328·0	286·3	0·871
5	6·18 } unreliable	249·9	240·0	0·960
6	7·42 }	157·8	155·0	0·982
7	8·6	105·2	105·2	1·000
8	9·8	52·6	52·6	1·000
9	11·0	52·6	52·6	1·000
10	12·0	39·5	39·3	1·000
11	...	26·3	26·3	1·000
12	...	39·5	39·5	1·000
13	...	19·8	19·8	1·000
14	...	19·8	19·8	1·000
15	...	6·5	6·6	1·000
16	...	13·2	13·2	1·000
17	...	13·2	13·2	1·000
18	...	13·2	6·5	1·000
19	...	6·5	3·3	1·000
20	...	3·3	6·6	1·000
21	...	6·6	13·2	1·000
22	...	13·2	3 3	1·000
23	...	3·3	1·3	1·000
24	...	1·3	6·6	1·000
25	...	6·6	1·3	1 000
26	...	1·3	0	1·000
27	...	0	3·3	1·000
28	...	3·3	0	1·000
29	...	0	0	1·000
30	...	0	1·3	1·000
31	...	1·3	1·3	1·000
32	24·0	1·3	1·3	1·000
33	25·0	1·3	...	1·000

The agreement of the numbers of the last column with those calculated by Silberstein's formula (with $N = 0{\cdot}572$ and $\sigma = 0{\cdot}0973$), as given in his paper, is manifestly a very pronounced one. The differences between the observed and calculated values are even in the case of the 3 grain clumps entirely within the limits of experimental error, particularly concerning the area measurements.

In continuance of this work the theoretical formula will be subject to further tests, the results of which will be published shortly.

Rochester, N.Y.
January 24, 1922.

XIX. *Quantum Theory of Photographic Exposure.* By L. Silberstein, *Ph.D.**

1. THE purpose of the present paper is to describe a first attempt at a light-quantum theory of photographic exposure, or of the production of the so-called latent image, with the immediate consequences of such a theory and some of its experimental tests.

The silver-halide grains of an emulsion spread over a plate or a film base may be considered (apart from the smallest grains) as small flat plates, of comparatively small thickness, which in a dry emulsion lie almost parallel to the base. The sizes a (areas) of these plates range from submicroscopic ones up to 18 or 20 square microns. Even the most uniform emulsions obtainable in practice consist of grains of different sizes, the distribution of sizes among the grains being in each case characterized by what is technically called "the frequency curve" of an emulsion. In what follows the number, per unit area of the photographic plate, of grains whose areas range from a to $a+da$ will be denoted by $f(a)da$, and a photographic emulsion will be shortly referred to as being of the type $f(a)$. For certain emulsions $f(a)$ is, with good approximation, an exponential, for others a Gaussian error-function of the area a, and so on.

2. Without, for the present, dwelling any longer upon details of this kind, we may pass at once to our main subject.

According to Einstein's well-known hypothesis of 1905 light does not consist in a continuous distribution of energy, as in the classical theory, but is entirely split up into *light quanta* or discrete parcels of very concentrated monochromatic light, each parcel containing a quantum of energy, $h\nu=h\lambda/c$, in obvious symbols. Somewhat more generally we may assume that only a fraction αE of the total light energy E is thus split into concentrated parcels, the remainder $E_0=(1-\alpha)E$ being distributed continuously †, without however prejudicing the possibility of E_0 being zero. Then, if E be the energetic value of a monochromatic

* Paper read December 28, 1921, at the Toronto meeting of the American Physical Society in affiliation with Section B of the American Association for the Advancement of Science. Communication No. 139 from the Research Laboratory of the Eastman Kodak Company.

† Somewhat as in E. Marx's theory of "concentration places" or of "light specks" as suggested by Sir J. J. Thomson.

"exposure" of wave-length λ, the number of light-quanta contained in it will be

$$n=\frac{\alpha E}{hc}\lambda. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

From a recent conversation with Einstein, there are weighty reasons for making $\alpha=1$. But since we do not prejudice its value, there is no harm in retaining this coefficient in the formula.

Now, let us assume that the necessary and sufficient condition for a silver-halide grain to be *affected*, i. e., to be made developable (entirely or in part) is that it should *absorb one light-quantum*.

Moreover, let us assume that a grain * does absorb a light quantum *whenever it is fully hit by one*, of a sufficiently high frequency ν_c or of a wave-length not exceeding a certain value λ_c.

There are perhaps some experimental hints or more or less good reasons for making these two assumptions, but we need not stop to consider them here. It will be time to reject or to modify them when they are contradicted by photographic experiments. Nor is it necessary to enter here upon the mechanism by means of which a silver-halide grain is affected by a light-quantum, whether it be the knocking out of an electron, as suggested by Joly, or something entirely different. For none of such details will influence our main argument, to be treated in the next section. Only when we come to consider the dependence of the photographic effect upon the wave-length will it be interesting to consider the photoelectric hypothesis and necessary to take account of the fact that a photo-electron is not liberated unless the frequency exceeds a certain, the so-called critical value. Under these circumstances λ_c appearing in our second assumption will stand for the critical wave-length as known from Photo-Electricity.

Again, whether a grain being "affected" is made developable in part only or throughout its whole area (no matter how large) is, in view of the kind of the contemplated experimental tests, of no great importance. As a matter of fact, however, there is good evidence that a grain is always made developable as a whole, no matter what its size, and this seems even to hold for "clumps" or aggregates of several smaller grains, as will be explained hereafter. If so, then our formulæ, to be developed presently, will give not merely the number of affected grains but also, by integration, the total "mass" made developable and hence also the photographic density. But we may as well remain content with the formulæ for the number (k) of grains affected, and count these in all experimental tests. This, far from being a disadvantage, will enable us to subject the proposed theory to more precise, though at the same time more severe tests.

One more remark. It will be understood that when we come to adopt the photo-electric hypothesis, a grain "affected" will stand for a grain

* Or perhaps, more generally, one of every p grains hit, where p is a number to be determined by experiment, but presumably equal to unity.

which as a whole has been deprived of even a single electron only. It need not lose more than one electron in order to be made entirely developable. According to Professor Joly's original hypothesis * the latent image "is built up of ionised atoms or molecules." In our connexion, this does not mean that for every pair of atoms, say Ag Br, there is one electron liberated. Since every grain of silver bromide (as well as of silver chloride) is a crystalline, to wit a simple cubic space-lattice arrangement of Ag and Br atoms †, we may as well consider the whole grain as a single molecule. Such a crystalline structure being hit by a light-quantum and deprived of but a single electron, may well become susceptible throughout to the subsequent action of a developer.

3. With the assumptions just made the question is reduced to a mere problem in probabilities.

Consider first the ideal case of equal grains. Let there be upon an area S of the photographic plate (say unit area) N grains. Let a be the size (area) of each of them divided by S, and let n light-quanta impinge upon S, due allowance having been made for those which may be reflected or absorbed by the gelatine. The problem consists in finding the number k of grains hit and (if $p=1$) affected by this light exposure.

Roughly, this simple problem can be treated as if N, k were continuous quantities, in the following way, familiar from many other instances ("mass-law"). At any stage the number of unaffected grains is $N-k$, representing an available fraction $a(N-k)$ of the total area S. Thus, if further dn light-quanta be thrown upon S, and if their transversal dimensions be negligible as compared with those of the grains ‡, the corresponding increment of k will be

$$dk=a(N-k)\,dn,$$

and since $k=0$ for $n=0$, this gives at once

$$k=N(1-e^{-na}), \quad . \quad . \quad . \quad . \quad . \quad . \quad (A)$$

which in fact will presently appear to be correct enough except for the practically unimportant cases of small N or small $N-k$.

More rigorously, but provided always that N at least is a large number §, the required formula can be obtained by the

* Nature, 1905, p. 308.

† R. B. Wilsey, Phil. Mag. xlii. p. 262 (1921).

‡ An assumption which will be given up in the sequel.

§ Which will be the case if S is taken large enough. Since plates and films in actual use contain as many as 10^7 grains per cm.^{2}, S can be made as small as one-hundredth mm.^{2}, and even less.

following reasoning. The total area of silver halide being a fraction Na of the area S of the plate, each of the nNa light-quanta will fall upon some grain. Of these nNa quanta one, say the first, will hit one grain, the next $\frac{N}{N-1}$ quanta will fall upon another grain, the next $\frac{N}{N-2}$ will hit yet another, and so on, up to $\frac{N}{N-k+1}$ quanta for the kth grain *. Thus the required relation between k and n will be

$$an=\frac{1}{N}+\frac{1}{N-1}+\ldots.+\frac{1}{N-k+2}+\frac{1}{N-k+1}.$$

Now by a well-known theorem of Analysis

$$1+\frac{1}{2}+\frac{1}{3}+\ldots.+\frac{1}{m}=C+\log m+\epsilon(m), \quad . \quad . \quad (2)$$

where C is Euler's constant and $0<\epsilon(m)\leq 1/m$. Thus in our case

$$\log\frac{N}{N-k}=an+\zeta,$$

where ζ lies between $-1/N$ and $1/(N-k)$ or practically, since N is at any rate a large number,

$$0\leq\zeta=\epsilon(N-k)<\frac{1}{N-k}.$$

Whence,

$$k=N(1-e^{-na-\zeta}). \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

In practicable experimental tests (counts of affected grains of a given, narrow size-class) the rôle of the correction ζ whose value $\zeta=\epsilon(N-k)$ can at any time be found by (2), may become perceptible only when the contemplated grain class is near exhaustion.

Thus, apart from such extreme cases, we have again, as in (A), the simple formula

$$k=N(1-e^{-na}), \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

which, though approximate only, will turn out to be accurate enough even for moderate values of N.

A thoroughly rigorous treatment of the probability problem valid for any numbers N, n, seems to be the

* It is scarcely necessary to say that statements such as "$N/N-1$ quanta hit another grain" are to be taken statistically as relating to averages over a group of many trials.

following one. Of the n quanta the total number falling each upon some of the N grains will be *

$$m = Na\, n,$$

Na being the total area of silver halide, with S as unit area. Thus the problem is reduced to finding the distribution of m quanta among N grains. Now, let $p_m(i)$ be the probability of affecting, in a single trial, a number i of the N grains, by the m quanta. By a combinatorial discussion which may be omitted here, I find

$$p_m(i) = \frac{N!\, a_{mi}}{N^m (N-i)!}, \quad \ldots\ldots \quad (5)$$

where a_{mi} may be most shortly described as identical with the number of ways in which a product of m different primes can be decomposed into i factors. These numbers, which will be known to many readers from combinatorial algebra, have the obvious properties

$$a_{m1} = a_{mm} = 1, \text{ for any } m,$$

$$a_{mi} = 0 \text{ for } i > m,$$

and satisfy the general recurrency formula

$$a_{mi} = a_{m-1,\,i-1} + i a_{m-1,\,i},$$

which enables us to write down successively without trouble any number of them.

Thus, up to $m = 10$, we have the following table which the reader may continue to extend at his leisure. Columns correspond to constant m, and rows to constant i.

1	1	1	1	1	1	1	1	1	1
0	1	3	7	15	31	63	127	255	511
	0	1	6	25	90	301	966	3025	9330
		0	1	10	65	350	1701	7770	34105
			0	1	15	140	2401	13706	76300
				0	1	21	266	3997	37688
					0	1	28	462	7231
						0	1	36	750
							0	1	45
								0	1

We may mention in passing that any a_{mi} can be represented by †

$$a_{mi} = \frac{1}{i!}\left[i^m - \binom{i}{1}(i-1)^m + \binom{i}{2}(i-2)^m - \ldots\ldots \pm \binom{i}{1}\right].$$

But for any numerical applications the table will be found more convenient.

---

* In a large number of trials of the same experiment.

† *Cf.* e. g. Lehrbuch d. Kombinatorik by E. Netto, Leipsig, 1901, p. 170.

By formula (5) we have for instance the probability of hitting but one grain

$$p_m(1)=\frac{1}{N^{m-1}},$$

which is an obvious result. The probability of hitting two grains will be

$$p_m(2)=(2^{m-1}-1)(N-1):N^{m-1},$$

which is (approximately) $2^m N$ times as large as the preceding probability, and so on. If m be kept constant, $p_m(i)$ will increase with growing i up to its largest value for $i=m$, if $m<N$, or for $i=N$, if $m>N$. But it would be futile to expect, on an average, that distribution to which corresponds the greatest probability *. For all other distributions have some generally non-negligible probabilities and these are by no means symmetrically spaced with respect to the largest one. The only reasonable way of determining the number k is to define it as the average of i taken over a large number M of trials. Out of these M trials a number M $p_m(i)$ † of trials will give each i grains affected, and the total number of grains affected in all M trials will be $\Sigma \mathrm{M} i p_m(i)$, to be summed from $i=1$ to $i=m$, if $m\leq N$ and to $i=N$ if $m>N$; but since $a_{mi}=0$ for $i>m$, we can as well extend the sum in each case from $i=1$ to $i=N$.

Dividing this sum by M we shall have the average number of grains hit in one trial, *i. e.*, by (5)

$$k=\frac{N!}{N^m}\sum_{i=1}^{N}\frac{i a_{mi}}{(N-i)!} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

This, with $m=Nan$, is the required rigorous formula for the number of grains affected, *i. e.*, hit once or more. In order to see how this complicated formula degenerates into (4), which, of course, will be our working formula, develop the sum in (6). Collecting the terms in N^{-1}, N^{-2} etc., and taking account of the values of a_{mi}, it will be found that

$$k=m-\frac{1}{N}\binom{m}{2}+\frac{1}{N^2}\binom{m}{3}-\ldots+\left(-\frac{1}{N}\right)^{m-1},$$

* If, say, $m<N$, the most probable distribution is the equipartition (a quantum per grain), corresponding to $p_m(m)=\frac{N!}{N^m(N-m)!}$. This would give as the number of grains affected $k=m=Nna$, or just the first term in the series development of (4), which would be hopelessly wrong unless mN were very small.

† With a deviation ruled by Bernoulli's law.

or, dividing by N and subtracting from unity,

$$1-\frac{k}{N}=1+\binom{m}{1}\left(-\frac{1}{N}\right)+\ldots+\binom{m}{m}\left(-\frac{1}{N}\right)^m,$$

i. e. ultimately

$$1-\frac{k}{N}=\left(1-\frac{1}{N}\right)^m. \quad . \quad . \quad . \quad . \quad . \quad (6')$$

This is rigorously equivalent to or identical with (6) for any m and N. Now, *for large* N, and any m, equation (6′) in which $(1-1/N)^N=1/e$, asymptotically, gives at once

$$\frac{k}{N}=1-e^{-\frac{m}{N}}=1-e^{-na}.$$

This is the connexion between the rigorous arithmetical formula and the exponential one.

It is needless to insist that under the conditions prevailing in all practicable experimental cases there is more than sufficient mathematical accuracy in formula (4)

$$k=N(1-e^{-na}),$$

which, apart from minor modifications, will henceforth be used in what follows.

This formula is of the familiar type proposed (1893) by Elder, with the notable difference, however, that while his exponent contained a free "parameter" or coefficient to be evaluated empirically and principally depending upon "grain-sensitivity" and wave-length, both coefficients in (4) are completely determined, and the exponent moreover shows an explicit and most essential dependence upon the size (a) of the grain, and in the right sense too, *i. e.*, giving an increase of the "speed" with grain size. The comparison with experimental facts of Elder's and of a number of other formulæ, constructed empirically, is too well-known to be discussed here *. Suffice it to say that, although it represents to a certain extent the photographic behaviour (the "characteristic" curve) of some emulsions, and particularly those with what is termed an "extended toe," it certainly shows considerable deviations from the observed characteristic curves. Yet it will not be forgotten that all these comparisons bore upon the resultant total photographic densities, containing or integrating the effects of grains belonging to a broad range of sizes (a), instead of equal

* *Cf.* for instance, a paper by Dr. F. E. Ross, Journ. Opt. Soc. Amer. vol. iv. p. 255 (1920).

grains, so that no better agreement could be expected. The refined experimental tests which are now in progress in this laboratory, and by means of which it is hoped to corroborate the proposed theory, deal, as they should, with *separate size-classes* of grains.

But questions concerning the comparison of the theory with experiment will be treated in a later part of the present and in subsequent papers.

4. Passing next to the case of an emulsion of any type $f(a)$, it can be easily proved that the approximate formula (4) will hold for each class of grains separately. In order to see this it is enough to consider the case of two distinct classes of grains. Thus, let there be N_1 grains of size a_1 and N_2 grains of size a_2 spread over the (unit) area S of the plate, and let n light-quanta be thrown upon S. Of these a number $m_1 = nN_1a_1$ will fall on the a_1-grains and a number $m_2 = nN_2a_2$ upon the a_2-grains. It remains only to be found how many a_1-grains will be hit by the m_1 quanta, and how many of the a_2-grains will be hit by the m_2 quanta. Now each of these is a problem of the kind we have already treated. The number k_1 of a_1-grains hit will be given by

$$\frac{1}{N_1} + \frac{1}{N_1 - 1} + \ldots + \frac{1}{N_1 - k_1 + 1} = na_1,$$

and similarly for the a_2-grains, so that k_1 and k_2 will each be determined by the previous formula for k with N, a replaced by N_1, a_1, and N_2, a_2 respectively. Similarly for an emulsion consisting of three or more classes of grains.

Thus, also, for an emulsion of any type $f(a)$, the number of grains ranging from a to $a + da$ hit and (if $p = 1$) affected by n impinging light-quanta will be $k_a\,da$, where, apart from the correction ζ

$$k_a = f(a)[1 - e^{-na}]. \quad . \quad . \quad . \quad . \quad . \quad (7)$$

The total area * of silver halide affected or made developable will be found by extending the integral

$$K = \int a k_a\,da \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

over the whole range of sizes, say from a_1 to a_2.

If, for instance, $f(a) = Ce^{-\mu a}$, say from $a = a_1$ to a_2, where C, μ are constants, as in the case of some films and plates

* It will be kept in mind that a stands for the "efficient" area of a grain (plate), *i.e.* for the orthogonal projection of the grain upon the film base.

investigated in this laboratory for their frequency curves, then, with A written for the total area of silver halide,

$$K=A-\frac{C}{(n+\mu)^2}\left\{[1+(n+\mu)a_1]e^{-(\mu+n)a_1}-[1+(n+\mu)a_2]e^{-(\mu+n)a_2}\right\}.$$

But this only by way of illustration. The fundamentally important thing will be formula (7), applicable to each size-class of grains separately. In fact the experimental verification of the theory now in progress in this laboratory deals, not so much with K but, as it should, microscopically, with $k=k_a$ for each class of grains separately, including clumps of grains.

Before passing to a further discussion and development of the elementary formula (7), but one more remark concerning the presence of more than one layer of grains. The case of two or more layers will at once be seen to be reducible to that of a single layer. In fact, either a grain of, say, the second layer and of size a is not shielded by any of the first layer or else it is thus screened off and only a part b of it remains uncovered. In the former case the grain in question will simply be classified among those of size a of the first layer, and in the latter case among those of size b. This will hold with respect to the exposure to the impinging light-quanta, and b will also be the contribution of the grain in question to the photographic density; for its covered part will remain inoperative. Similarly for three or more layers. In fine, the presence of a plurality of layers of grains will modify only the frequency curve $N_a=f(a)$ which would otherwise belong to a single layer. We shall henceforth assume that this factor has already been taken into account in constructing the function $f(a)$ or in microscopic counts of the grains within every particular size-class. We disregard here, of course, such factors as a possible absorption of light in additional strata of gelatine.

5. *Dependence upon wave-length.*—Once more return to the elementary formula (7) or (4). Denote by s the exponent so that

$$\frac{k}{N}=1-e^{-s}.$$

Under the more or less implicit assumption that the transversal dimensions of a light-quantum are negligible in comparison to those of a grain, we had $s=na$. But it will

be seen presently that such an assumption is too narrow and unnecessarily so.

In fact, substituting the number n of light-quanta from equation (1), the exponent s * will become $s=\frac{\alpha}{hc}E\lambda$, or, if we put for brevity

$$\beta=\alpha/hc, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

which may be considered as a constant,

$$s=\beta E\,.\,a\lambda. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Here E is the incident light energy (exposure) and λ the wave-length of the light assumed to be monochromatic. Thus the sensitivity exponent would be directly proportional to the wave-length, and the number k would, for constant E and a, increase steadily with the wave-length of the incident light up to the photoelectric critical value λ_c and then drop suddenly to zero,

$$\frac{k}{N}=1-e^{-\text{const.}\lambda},\ \lambda\leq\lambda_c,$$

$$k=0, \qquad \lambda>\lambda_c.$$

Now, such a sensitivity curve does not seem to resemble the familiar experimental sensitivity curves which show a more or less gentle maximum followed by a gradual decrease down to zero. It is true that such experimental curves † represent the resultant effect due to grains of a whole range of sizes, so that the $k-\lambda$ curve belonging to a single class of (equal) grains may well be of the said abrupt type,—a question to be decided only by micro-spectrographic experiments and counts now in progress. Yet it seems advisable even at this stage to provide for the possibility of smooth maxima preceding the critical wave-length λ_c.

This might be obtained by attributing to grains of different sizes different values of λ_c. For then, although the curve of each grain class would end abruptly, the superposition of such curves ending over a range of different abscissæ, might properly displace and smooth out the resultant maximum. The correctness of such an assumption (λ_c a function of a) can at any rate be tested by direct experiments ‡.

* s/E can be referred to as the "sensitivity exponent."

† Apart from the fact that they are not taken for $E=$const.

‡ Preparations for such experiments are now being made in this laboratory.

Another way is to take account of the possibly *finite transversal dimensions* of the light parcels, which may perhaps be comparable with those of the lesser silver halide grains.

Let us assume, therefore, that a light-quantum, of sufficiently high frequency, becomes efficient in affecting a grain only if it strikes it fully, or almost so. To fix the ideas, let r be the equivalent radius of a grain, *i. e.*, such that

$$a=\pi r^2,$$

and similarly let ρ be the average radius of the transversal section of a light parcel (so to call the space occupied by a quantum of energy). Then the efficient area of a grain to be substituted instead of a, will be

$$a'=\pi(r-\rho)^2=a\left[1-\frac{\rho}{r}\right]^2,$$

and we shall have for the exponent s, instead of (10),

$$s=na'=\beta Ea\left[1-\frac{\rho}{r}\right]^2\lambda, \text{ for } \rho\leq r,\ \lambda\leq\lambda_c. \quad (10\,a)$$

and $s=0$ for $\rho>r$ or $\lambda>\lambda_c$.

It remains to assume in a general way that ρ, which may be the average of section-radii different even for parcels of the same wave-length, is itself a function of the wave-length increasing with λ, without prejudicing, however, the particular form of this function. Certain easily ascertainable broad features of such a function and thence also of the resulting factor in s,

$$\phi(\lambda)=\lambda\left[1-\frac{\rho(\lambda)}{r}\right]^2, \quad . \quad . \quad . \quad . \quad . \quad (11)$$

will suffice to ensure a maximum of the sensitivity exponent between $\lambda=0$ and $\lambda=\lambda_c$. The value of λ_c itself may still turn out to depend on the size a of the grain and on its physical conditions as well. Every process which will make the liberation of a photoelectron from the grain (crystal) easier will lengthen λ_c. Part of the effect of sensitizing may arise in this manner. But questions of this kind must necessarily be postponed until some further experimental data are gathered. Of such a kind is also the question whether ρ (which, for a given λ, may also extend over a whole range of values) attains at all the semi-diameter r of even the smallest of the actual grains, and whether the corresponding wave-length λ_0 entailing the vanishing of s, exceeds or is smaller than λ_c as derived from direct photoelectric experiments. In absence of all knowledge concerning

the spatial properties of light-quanta it would be utterly unjustified either to deny or to assert that their lateral dimensions are at all comparable with those of a silver halide grain (of the order of one-tenth up to several microns) *. If, by way of illustration only, ρ is proportional to a power of λ, say $\rho = b\lambda^{\kappa}$, the only condition for the existence of a maximum of $\phi(\lambda)$, and therefore of sensitivity, will easily be found to be $\kappa > 0$. If this be satisfied, the maximum will occur at a wave-length λ_m given by

$$(2\kappa+1)b\lambda_m^{\kappa} = r,$$

increasing with the diameter of the grain and bearing to λ_0 the fixed ratio

$$\frac{\lambda_m}{\lambda_0} = (2\kappa+1)^{1/\kappa}.$$

As a matter of fact, the maximum sensitivity is known to shift (by two or three hundred A.U.) towards the red by making the grain coarser. But thus far too little is known of the quantitative aspect of such an effect to entitle one to consider the above equation as anything more than an illustrative example. The precise form of the function $\rho = \rho(\lambda)$ can only be derived from spectrographic experiments followed by microscopic grain counts, or if arrived at by a guess, has to be verified by them. Such experiments are now in preparation in this laboratory, and their results will be reported in due time. A shift of the maximum sensitivity towards the red or the infra-red can, of course, be brought about by a function form more general than a mere positive power of the wave-length.

6. *Generalities, and preliminary account of experimental tests.*—The chief and most immediate consequence of the proposed theory is the essential dependence of the proportionate number of grains affected, k/N, on the size a of the grain, viz., the rapid increase of the former and, therefore, of "the speed" of an emulsion with the latter. Now, it has been known for a long time that (*cæteris paribus*) the speed increases notably with the size of the grain, and we shall see from the experiments to be described presently to how

* According to E. Marx, *Annalen der Physik*, xli. pp. 161-190 (1913), the volume of a light-parcel, which according to him is only a "concentration place" within a continuous distribution of energy, is proportional to λ^4 and amounts for D-light to almost 8.10^{-7} cm.^{3}, which even with a length of 10 cm. (200,000 D-waves) would still give a section area 8.10^{-8} cm.^{2}, just of the order of about the average grain area. There is, of course, nothing cogent about Marx's estimate, yet the matter is not without interest.

large an extent this is actually the case. But perhaps the most tangible proof of the essential correctness of the assumption of spatially *discrete* as against continuous action*, seems to be the mere fact, disclosed by microscopic counts, that out of a number of apparently equal grains subjected to a sufficiently weak exposure one or two are affected while the others, nay their next neighbours, remain perfectly intact. It would be in vain to ascribe to these survivors a greater immunity or indifference to light. For it is enough to protract the exposure a little to make them succumb in their turn. Now such a behaviour is most typical of rain as contrasted with flood action, and the discrete light-quanta, hitting now this and now that grain, appear to be a most natural inference, while all attempts to bring into play the individual "sensitiveness" of the units seem to involve considerable difficulties.

As to the dependence of the number of grains affected upon the wave-length, little more of interest in the present connexion is known than the qualitative fact of a shift towards the red of the maximum sensitivity with increasing size of the grain. Moreover, the available curves representing the sensitivity across the spectrum concern the emulsion as a whole and not the separate a-classes of grain with which we are primarily concerned. Spectrographic and spectrophotometric experiments of such a kind, to be aided by direct photo-electric measurements, are now in progress in this laboratory, and all discussions involving wave-length will best be postponed until the results of these experiments and of laborious microscopic counts are forthcoming.

Before passing to the mentioned quantitative test of the dependence on size, but one more general remark. The reader will have noticed the complete absence of the time-variable in all our formulæ, the exposure entering only through the total number n of light-quanta or through the energy E which, in obvious symbols, is $\int I\,dt$. The proposed theory, therefore, as thus far developed, does not take any account of the little infringements against the reciprocity law †, in short, of the so-called "failure of the reciprocity law." Now, it is by no means my intention to deny the

* A rain as against a flood, of light, that is.

† This early law asserted the dependence of the photographic effect (density) upon I and t only through the total incident energy or exposure $\int I dt$. For constant intensity this is It, whence the name of the law, relating to intensity and exposure-time as factors of a constant product.

reality of these infringements which have been extensively studied and condensed into empirical formulæ by Abney, Schwarzschild, Kron and others. But it has seemed inadvisable to encumber the very beginnings of the proposed theory by complicated details of such a kind*.

The failure of the reciprocity law can more profitably be taken up later on, after the fundamentals of the theory have been somewhat solidified and extensively tested, and the prospects of mastering the "failure" theoretically are by no means averse, a very promising scheme seeming to lie in the possibility (suggested by Joly and taken up by H. S. Allen) of the liberated photo-electrons being regained by some of the grains which were deprived of them by previous impacts of light-quanta. In fine, the failure of the reciprocity law as well as the facts known under the head of "reversal" have at first to be neglected and considered as future problems for the light-quantum theory united with Joly's photo-electric theory, problems to which these combined theories seem well equal.

To pass to numerical facts, a short description will now be given of the results of certain experiments undertaken in this laboratory by A. P. H. Trivelli and Lester Righter† which seem to corroborate the proposed theory most emphatically. In order to have a much wider range of sizes a than is usually afforded by the single grains, Trivelli and Righter applied their counts and area measurements to clumps of from one to as many as 33 grains, basing themselves upon the well-supported assumption that if one of the component grains be affected, the whole clump is made developable. (This, at any rate, is the behaviour

* R. E. Slade and G. I. Higson, "Photochemical Investigations of the Photographic Plate," Proc. Roy. Soc. xcviii. pp. 154–170 (1920), on the contrary, make the failure of the reciprocity law their point of departure. They mention at the very beginning (p. 156) the possibility of a light-quantum theory and write the Elder-type of formula in I, t, remarking even that its coefficient would have a different value for each size of grain, but being discouraged by or rather preoccupied with the failure of the reciprocity law do not enter into the details of the probability problem, which would have disclosed them the structure of that coefficient, and without much ado dismiss the quantum theory as "impossible." Independently of Slade and Higson the possibility of a discrete theory (radiation in "filaments") is mentioned by F. E. Ross, Astrophys. Journ. vol. lii. p. 95 (1920). Dr. Ross, without being prejudiced against such a theory, notes even that it would lead rationally to a mass-law, but does not enter into the details of the probability problem and does not develop the theory.

† For technical details of these laborious experiments, see Trivelli and Righter's own note in this issue of the Phil. Mag. p. 252.

of the larger, flat grains piled upon each other in part, although the smaller, spherical grains, in less intimate contact, may perhaps behave differently.) Such being the case, their experimental results should be covered by our formula with a written for the area of the *whole clump*, no matter how large and how numerous its components. This has seemed a rather severe test but the more so tempting and instructive. Since all the classes of clumps were given, in each trial, a unique exposure (through a blue filter specified *loc. cit.*) and there was no question of varying λ, it will be most convenient to retain in the corresponding formula the original light-quantum number n as the parameter common to all clumps. Thus the formula to be tested becomes

$$\frac{k}{N}=1-e^{-s}, \quad s=na'=na\left(1-\frac{\rho}{r}\right)^2,$$

or somewhat more conveniently for computations, if $\sigma=\pi\rho^2$ be the (average) area of the transversal section of a light parcel,

$$\log\left(1-\frac{k}{N}\right)=-na\left[1-\sqrt{\frac{\sigma}{a}}\right]^2. \quad . \quad . \quad (12)$$

In the following table the first column gives the number of grains in a clump, the second the average area a of a clump in square microns, and the third column the percentage of clumps affected out of all (N) clumps of each kind originally present, *i. e.*

$$y=\frac{100\,k}{N},$$

as deduced by Trivelli and Righter from their observations.

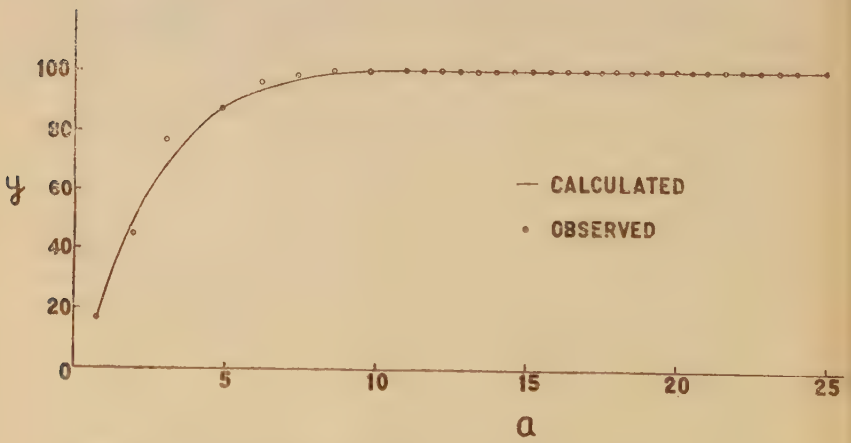
Clumps of	a in μ^2.	$y_{obs.}$	$y_{calc.}$	Δy.
1 grain	0·75⁴	16·5	16·2	+0·3
2 grains	1·92⁵	44·9	48·4	−3·5
3 „	3·03	76·6	68·9	+8·3
4 „	4·88	87·1	87·3	−0·2
5 „	6·1⁸	96·0	93·3	2·7
6 „	7·4²	98·2	96·4	1·8
7 „	(8·6)	100	98·0	2·0
8 „	(9·8)	100	99·6	0·4
9 „	(11·)	100	99·8	0·2
10 „	(12·)	100	100	0·0
etc., etc.	etc., etc.	idem.	idem.	idem.
32 grains	>24	100	100	0·0
33 „	>25	100	100	0·0

The most reliable a-values are those for the clumps of one and of two grains, being averages of the largest numbers

of individual clumps ; the following a's are gradually less reliable ; from the 7-grain clumps onward the areas (bracketed) are only extrapolated, but since here y has practically reached 100, no greater accuracy is needed. The fourth column contains the theoretical values of y following from formula (12) with the constants, determined from observations 1, 2, 4,

$$\left.\begin{aligned} n &= 0{\cdot}572^4 \text{ per } \mu^2, \\ \sigma &= 0{\cdot}097^3\ \mu^2, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (12\,a)$$

or $\rho = 0{\cdot}176\mu$. The fifth column gives $\Delta y = y_{\text{obs.}} - y_{\text{calc.}}$. The agreement is certainly very pronounced, the differences being, perhaps with the exception of the third, well within the limits of experimental error chiefly in the a-estimates. The fitting could be made even closer by retouching the constants n, σ, but this is scarcely worth while, the formula itself being of a statistical nature, and the agreement being good enough as it stands. The same is manifest also from the figure giving a graphical representation of the last columns of the table.



The reader might think that the finite section σ, or whatever this parameter may stand for, has been forced upon the light-quantum and that the observations might perhaps be as well represented with $\sigma = 0$, and another value of n. But actually, just the contrary has been the case, inasmuch as the author first tried the simpler formula $\log(1 - k/\Lambda) = -na$, and then only found himself compelled to take in the correction factor as given in (12). In fact,

dividing the observed values of $\log(1-k/N)$ by the areas the reader will find that the quotient increases considerably and systematically, apart from a casual drop at the fourth clump, throughout the whole series of the clumps. Thus, the correction factor seems to come in quite spontaneously.

On the other hand, there is nothing unlikely about the values of either of the constants ($12\,a$). Our units of area being here square microns, we should have 57 millions light-quanta per cm.^{2} (about which judgment has to be suspended until absolute energy measurements are available), and as the cross section of the space occupied by each of them (on an average) a little less than *one-tenth* of a square micron or a diameter of about 0·35 micron. Since each is presumably of about the order of a million wave-length long, they are rather slender at that cross section, and, instead of light parcels, as they were called above, would perhaps deserve rather the name of *light darts*. In Einstein's own theory there is nothing on which to base an estimate of the volume occupied by a light-quantum, but on Marx's less radical views this is about 8.10^{-7} cm.^{3} for D-light and proportional to λ^4, and therefore in our case (narrow blue spectrum region with maximum at $\lambda = 0{\cdot}470\mu$) about 3.10^{-7} cm.^{3}, which with the said cross section would mean a length of $3.10^6\mu$ or over six million wave-lengths. But this by the way only. The important thing is to see whether the above numerical value of the average cross section of blue light darts will continue to fulfil its function with regard to the remaining "steps" (weaker and weaker exposures) of plates coated with the same emulsion, the above being the highest "step." These have just been completed in this laboratory, *cæteris paribus* with the above one, and are now being subjected to counts and area measurements. This material will also serve to test the constancy of k/N if, varying n and a', their product is kept constant.

An account of the results of these and of several other experiments now in progress will be given in future papers.

I gladly take this opportunity to express my best thanks to Dr. C. E. K. Mees for having proposed to me the problem of "discriminating, if possible, between the consequences of a discrete and a continuous exposure theory," and to my colleagues Trivelli and Righter for furnishing the results of their experiments.

Rochester, N.Y.,
January 19, 1922.

XX. *An Analytical Discrimination of Elastic Stresses in an Isotropic Body.* *By* R. F. Gwyther, *M.A.**

Sir G. B. Airy obtained from mechanical considerations (British Association Reports, Cambridge, 1862) a solution in Cartesian coordinates of the mechanical stress-equations, but he ignored all elastic requirements. In one sense this paper may be regarded as an extension of Airy's scheme, though it has nothing in common with that scheme either in general plan or detail.

The method is purely analytical, depending upon general solutions of the mechanical stress-equations and upon the development of a scheme for the selection of those stress-systems which satisfy the stress-strain relations, briefly called Hooke's law, from the general mechanical stress-systems.

It is shown that the elements of a mechanical stress depend upon an arbitrary primary stress-system, and, to form a connexion with the stress-strain relations, I introduce a subsidiary, but allied, stress-system which is such that the vector system naturally deduced from it possesses inherent qualities distinctive of the displacement corresponding to an elastic stress-system.

The main body of the paper consists in discussing the requirements necessary to ensure that the stress-system should be an elastic stress-system.

The displacement, according to this method, becomes somewhat incidental, however necessary, and the elements of stress are given prominence. There are no displacement equations.

In the first instance I deal with a body under tractions only, and extend the scope of the results later.

I. Introduction.

1. By first treating certain ancillary matters as lemmas, the steps in the final stages can progress more steadily.

The mechanical equations of stress in a body under

* Communicated by the Author.

tractions only are, in Cartesian coordinates,

$$\frac{\partial P}{\partial x}+\frac{\partial U}{\partial y}+\frac{\partial T}{\partial z}=0,$$
$$\frac{\partial U}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial S}{\partial z}=0,$$
$$\frac{\partial T}{\partial x}+\frac{\partial S}{\partial y}+\frac{\partial R}{\partial z}=0; \quad \ldots\ldots \quad (1)$$

and they are identically satisfied by the values

$$P = -\frac{\partial^2\theta_2}{\partial z^2}-\frac{\partial^2\theta_3}{\partial y^2}+2\frac{\partial^2\psi_1}{\partial y\partial z},$$
$$Q = -\frac{\partial^2\theta_1}{\partial z^2}-\frac{\partial^2\theta_3}{\partial x^2}+2\frac{\partial^2\psi_2}{\partial x\partial z},$$
$$R = -\frac{\partial^2\theta_1}{\partial y^2}-\frac{\partial^2\theta_2}{\partial x^2}+2\frac{\partial^2\psi_3}{\partial x\partial y},$$
$$S = \frac{\partial^2\theta_1}{\partial y\partial z}+\frac{\partial^2\psi_1}{\partial x^2}-\frac{\partial^2\psi_2}{\partial x\partial y}-\frac{\partial^2\psi_3}{\partial x\partial z},$$
$$T = \frac{\partial^2\theta_2}{\partial x\partial z}-\frac{\partial^2\psi_1}{\partial x\partial z}+\frac{\partial^2\psi_2}{\partial y^2}-\frac{\partial^2\psi_3}{\partial y\partial z},$$
$$U = \frac{\partial^2\theta_3}{\partial x\partial y}-\frac{\partial^2\psi_1}{\partial x\partial z}-\frac{\partial^2\psi_2}{\partial y\partial z}+\frac{\partial^2\psi_3}{\partial z^2}. \quad \ldots \quad (2)$$

These contain six arbitrary and general functions, and form the general solution of (1).

Lemma A.

These six arbitrary functions have the same mode of resolution on transformation of axes as the elements of stress.

For proof I use the method employed in the subject of differential-invariants.

Imagine the axes of coordinates to be rotated about their own positions by the infinitesimal amounts ω_x, ω_y, ω_z, and consider the consequent changes in whatever quantities we may be considering—components of a vector, elements of a stress, etc. These changes will be linear functions of ω_x, ω_y, ω_z.

For example, in the case of the components of a vector u, v, w, the changes will be :

$$\text{in } u, \quad -\omega_y w + \omega_z v,$$

$$\text{in } v, \quad -\omega_z u + \omega_x w,$$

$$\text{in } w, \quad -\omega_x v + \omega_y u.$$

These changes will now be represented by partial differential-operators Ω_1, Ω_2, Ω_3 arranged to produce the coefficients of ω_x, ω_y, ω_z in the quantities considered respectively.

Thus, for vectors generally,

$$\Omega_1 = w\frac{\partial}{\partial v} - v\frac{\partial}{\partial w},$$

$$\Omega_2 = u\frac{\partial}{\partial w} - w\frac{\partial}{\partial u},$$

$$\Omega_3 = v\frac{\partial}{\partial u} - u\frac{\partial}{\partial v}.$$

This covers direction-cosines (l, m, n), and may be made the basis of most operators. For example, we may deduce for stresses generally :

$$\Omega_1 = 2S\left(\frac{\partial}{\partial Q} - \frac{\partial}{\partial R}\right) + (R-Q)\frac{\partial}{\partial S} - U\frac{\partial}{\partial T} + T\frac{\partial}{\partial U},$$

$$\Omega_2 = 2T\left(\frac{\partial}{\partial R} - \frac{\partial}{\partial P}\right) + (P-R)\frac{\partial}{\partial T} - S\frac{\partial}{\partial U} + U\frac{\partial}{\partial S},$$

$$\Omega_3 = 2U\left(\frac{\partial}{\partial P} - \frac{\partial}{\partial Q}\right) + (Q-P)\frac{\partial}{\partial U} - T\frac{\partial}{\partial S} + S\frac{\partial}{\partial T}.$$

The differential coefficients $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ resolve as components of a vector, but, for simplicity, we must introduce a different notation.

Write d_1 for $\partial/\partial x$, d_2 for $\partial/\partial y$, d_3 for $\partial/\partial z$, and d_{11} for $\partial^2/\partial x^2$, d_{12} for $\partial^2/\partial x \partial y$, d_{22} for $\partial^2/\partial y^2$, and so on ; then for first differential coefficients

$$\Omega_1 = d_3\frac{\partial}{\partial d_2} - d_2\frac{\partial}{\partial d_3}, \text{ etc.,}$$

as for vectors, and we deduce for second differential coefficients

$$\Omega_1 = d_{13}\frac{\partial}{\partial d_{12}} - d_{12}\frac{\partial}{\partial d_{13}} + 2d_{23}\left(\frac{\partial}{\partial d_{22}} - \frac{\partial}{\partial d_{33}}\right) - (d_{22} - d_{33})\frac{\partial}{\partial d_{23}},$$

$$\Omega_2 = d_{12}\frac{\partial}{\partial d_{23}} - d_{23}\frac{\partial}{\partial d_{12}} + 2d_{13}\left(\frac{\partial}{\partial d_{33}} - \frac{\partial}{\partial d_{11}}\right) - (d_{33} - d_{11})\frac{\partial}{\partial d_{13}},$$

$$\Omega_3 = d_{23}\frac{\partial}{\partial d_{13}} - d_{13}\frac{\partial}{\partial d_{23}} + 2d_{12}\left(\frac{\partial}{\partial d_{11}} - \frac{\partial}{\partial d_{22}}\right) - (d_{11} - d_{22})\frac{\partial}{\partial d_{12}}.$$

It is required to establish that in (2) $\theta_1, \theta_2, \theta_3, \psi_1, \psi_2, \psi_3$ act as stresses. Actually I shall assume that the operators $\Omega_1, \Omega_2, \Omega_3$ act upon them as upon the elements of stress, and first examine this hypothesis when the three operators separately act on the six equalities in (2).

Selecting the first equality in (2) and applying the operator Ω_1 on each of the two sides of the equality, then

$$\Omega_1 P = 0;$$

and

$$-\Omega_1\frac{\partial^2\theta_2}{\partial z^2} - \Omega_1\frac{\partial^2\theta_3}{\partial y^2} + 2\Omega_1\frac{\partial^2\psi_1}{\partial y\,\partial z}$$

becomes

$$-2\frac{\partial^2\psi_1}{\partial z^2} + 2\frac{\partial^2\psi_1}{\partial y^2} + 2\frac{\partial^2(\theta_3-\theta_2)}{\partial y\,\partial z} + 2\frac{\partial^2(\theta_2-\theta_3)}{\partial y\,\partial z} - 2\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right)\psi_1,$$

which is null, and the hypothesis is not negated.

In fact, the hypothesis is not negated in any case.

The argument is then as follows :—If we had written $\Omega_1\theta$, $\Omega_2\theta$, etc. in the equations, and had proceeded to find these 18 quantities, we should have 18 linear equations from which to find them. The solution is therefore unique, and cannot differ from that employed by hypothesis and found to satisfy the 18 tests.

We shall therefore regard $\theta_1, \theta_2, \theta_3, \psi_1, \psi_2, \psi_3$ as acting as elements of a stress, though they have not the proper dimensions.

It is proposed to find the conditions which must exist between these primary arbitrary stresses in order that the elements P, Q, R, S, T, U may be elements of an elastic stress.

Note.—If, in the ordinary notation for strains, we give a, b, c each one-half of the usual value given to it, strains would follow the same laws of composition and resolution as stresses, and would therefore have the same differential operators. In this paper, I shall use a, b, c in this sense—that is, one-half of their usual value.

Lemma B.

Except when we have reasons for keeping the expressions quite general, it will suffice to limit the arbitrary stress-system to such stresses as have the co-ordinate axes as their principal axes.

If in equation (2) the elements of stress, P, Q, etc., are made zero, the set of equations will then be recognized as indicating that $\theta_1=e$, $\theta_2=f$, $\theta_3=g$, $\psi_1=a$, $\psi_2=b$, $\psi_3=c$, where e, f, g, etc. are elements of strain arising from some arbitrary displacement.

Hence on the right-hand side of equations (2) we may always replace θ_1 by θ_1-e, θ_2 by θ_2-f, θ_3 by θ_3-g, ψ_1 by ψ_1-a, ψ_2 by ψ_2-b, ψ_3 by ψ_3-c. Consequently we may eliminate several sets of three functions, such as ψ_1, ψ_2, and ψ_3, when some displacement is possible which makes, say, $\psi_1=a$, $\psi_2=b$, $\psi_3=c$.

Hence

$$\mathrm{P}=-\frac{\partial^2\theta_3}{\partial z^2}-\frac{\partial^2\theta_3}{\partial y^2}, \text{ etc.},$$

$$\mathrm{S}=\frac{\partial^2\theta_1}{\partial y\,\partial z}, \text{ etc.},$$

which is the form given by Airy, is a quite general form of solution, although for the purpose of this paper the full form given in (2) is requisite until we have decided upon some particular set of axes.

2. The choice of a vector to represent the displacement, and the descriptive criterion of elastic stress.

The mechanical stress has been represented in terms of an arbitrary stress-system, and it is possible and desirable to represent the displacement in terms of a similar stress-system.

For this purpose I form a subsidiary stress-system,

indicated by

$$\frac{\theta_1+\theta_2+\theta_3}{2}-\theta_1,\quad \frac{\theta_1+\theta_2+\theta_3}{2}-\theta_2,\quad \frac{\theta_1+\theta_2+\theta_3}{2}-\theta_3,$$

$$-\psi_1,\quad -\psi_2,\quad -\psi_3.$$

This subsidiary system may be described as complementary to the primary stress-system, in the sense that the two together form a hydrostatic pressure whose intensity is one-half of the sum of the principal stresses or one-half of the First Invariant of the primary stress-system.

I shall form the assumed components of the displacement from the elements of this subsidiary stress in the manner of forming a force-system from a stress-system.

Thus, I shall write

$$\begin{aligned}
2nu &= \frac{\partial}{\partial x}(\theta_2+\theta_3-\theta_1)-2\frac{\partial\psi_3}{\partial y}-2\frac{\partial\psi_2}{\partial z},\\
2nv &= -2\frac{\partial\psi_3}{\partial x}+\frac{\partial}{\partial y}(\theta_3+\theta_1-\theta_2)-2\frac{\partial\psi_1}{\partial z},\\
2nw &= -2\frac{\partial\psi_2}{\partial x}-2\frac{\partial\psi_1}{\partial y}+\frac{\partial}{\partial z}(\theta_1+\theta_2-\theta_3). \quad . \quad . \quad (3)
\end{aligned}$$

Forming the values of S, T, U from these on the elastic stress-strain hypothesis we have

$$S = \frac{\partial^2\theta_1}{\partial y\,\partial z}-\frac{\partial^2\psi_2}{\partial x\,\partial y}-\frac{\partial^2\psi_3}{\partial x\,\partial z}-\frac{\partial^2\psi_1}{\partial y^2}-\frac{\partial^2\psi_1}{\partial z^2},$$

etc.

On equating these to the values for the same elements given in (2), we find they require

$$\nabla^2\psi_1 = 0,\quad \nabla^2\psi_2 = 0,\quad \nabla^2\psi_3 = 0. \quad . \quad . \quad (4)$$

Since $\{\theta_1, \theta_2, \theta_3, \psi_1, \psi_2, \psi_3\}$ act on transformation of coordinates as elements of stress, it follows that the system must consist of a hydrostatic pressure and a general stress-system, each of the elements of which is a Spherical Harmonic. That is,

$$\theta_1 = \phi+\chi_1,\quad \theta_2 = \phi+\chi_2,\quad \theta_3 = \phi+\chi_3,$$

where

$$\nabla^2\chi_1 = 0,\quad \nabla^2\chi_2 = 0,\quad \nabla^2\chi_3 = 0. \quad . \quad . \quad (5)$$

This is the descriptive criterion of an elastic stress-system.

3. Completion of the discrimination. The metric criterion.

The remaining requirements of the stress-strain relations may be written

$$P-Q = 2n\left(\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}\right),$$

$$Q-R = 2n\left(\frac{\partial v}{\partial y}-\frac{\partial w}{\partial z}\right),$$

$$P+Q+R = (3m-n)\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right). \quad . \quad . \quad . \quad (6)$$

which are to be completed from the values in (2) and (3). The first two only confirm the descriptive criteria of (4) and (5). The last leads to

$$-\nabla^2(\theta_1+\theta_2+\theta_3)+\left\{\frac{\partial^2\theta_1}{\partial x^2}+\frac{\partial^2\theta_2}{\partial y^2}+\frac{\partial^2\theta_3}{\partial z^2}\right.$$
$$\left.+2\frac{\partial^2\psi_1}{\partial y\,\partial z}+2\frac{\partial^2\psi_2}{\partial x\,\partial z}+2\frac{\partial^2\psi_3}{\partial x\,\partial y}\right\}$$
$$=\frac{3m-n}{2n}\left[\nabla^2(\theta_1+\theta_2+\theta_3)-2\left\{\frac{\partial^2\theta_1}{\partial x^2}+\frac{\partial^2\theta_2}{\partial y^2}+\frac{\partial^2\theta_3}{\partial z^2}\right.\right.$$
$$\left.\left.+2\frac{\partial^2\psi_1}{\partial y\,\partial z}+2\frac{\partial^2\psi_2}{\partial x\,\partial z}+2\frac{\partial^2\psi_3}{\partial x\,\partial y}\right\}\right],$$

and therefore to

$$(3m+n)\nabla^2(\theta_1+\theta_2+\theta_3)$$
$$=6m\left\{\frac{\partial^2\theta_1}{\partial x^2}+\frac{\partial^2\theta_2}{\partial y^2}+\frac{\partial^2\theta_3}{\partial z^2}+2\frac{\partial^2\psi_1}{\partial y\,\partial z}+2\frac{\partial^2\psi_2}{\partial x\,\partial z}+2\frac{\partial^2\psi_3}{\partial x\,\partial y}\right\}$$

or

$$(m+n)\nabla^2\phi$$
$$=2m\left\{\frac{\partial^2\chi_1}{\partial x^2}+\frac{\partial^2\chi_2}{\partial y^2}+\frac{\partial^2\chi_3}{\partial z^2}+2\frac{\partial^2\psi_1}{\partial y\,\partial z}+2\frac{\partial^2\psi_2}{\partial x\,\partial z}+2\frac{\partial^2\psi_3}{\partial x\,\partial y}\right\},$$
$$. \quad . \quad . \quad (7)$$

which is the metric criterion and completes the discrimination sought for.

This can be integrated, and gives ϕ in terms of the χ's and ψ's, of which we can always take the ψ's to be null, when desirable.

This completes the investigation for Cartesian coordinates under normal tractions only and with no inertia terms. With the conditions in (4), (5), and (6), the equations (2) give the general elastic stress-system, and these conditions discriminate an elastic stress-system from any other mechanical stress-system.

4. The inclusion of inertia terms.

We must now modify equations (1) by writing $\rho\ddot{u}$, $\rho\ddot{v}$, $\rho\ddot{w}$ on the right-hand side, where u, v, w are to have the values given in (3).

We consequently replace P by $P-\rho(\ddot{\theta}_2+\ddot{\theta}_3-\ddot{\theta}_1)/2n$, Q by $Q-\rho(\ddot{\theta}_3+\ddot{\theta}_1-\ddot{\theta}_2)/2n$, R by $R-\rho(\ddot{\theta}_1+\ddot{\theta}_2-\ddot{\theta}_3)/2n$, S by $S+\rho\ddot{\psi}_1/n$, T by $T+\rho\ddot{\psi}_2/n$, U by $U+\rho\ddot{\psi}_3/n$.

With these alterations the equations (2) still hold good.

In forming our criteria, we equate values found from (2) to values given by the stress-strain relations deduced from (3). In these latter P, Q, R, S, T, U are to have the original values of these quantities and not those which replace them as above.

Consequently, as our first step in the criteria, in place of $\nabla^2\psi_1=0$ etc. we obtain

$$n\nabla^2\psi_1=\rho\ddot{\psi}_1,\quad n\nabla^2\psi_2=\rho\ddot{\psi}_2,\quad n\nabla^2\psi_3=\rho\ddot{\psi}_3,$$

and similarly

$$n\nabla^2\chi_1=\rho\ddot{\chi}_1,\quad n\nabla^2\chi_2=\rho\ddot{\chi}_2,\quad n\nabla^2\chi_3=\rho\ddot{\chi}_3. \qquad (8)$$

In place of the last stage which gave the metric criterion, we find

$$\frac{\rho}{2n}(\ddot{\theta}_1+\ddot{\theta}_2+\ddot{\theta}_3)-\nabla^2(\theta_1+\theta_2+\theta_3)+\left\{\frac{\partial^2\theta_1}{\partial x^2}+\frac{\partial^2\theta_2}{dy^2}+\frac{\partial^2\theta_3}{dz^2}+2\frac{\partial^2\chi_1}{\partial y\,\partial z}+2\frac{\partial^2\chi_2}{\partial x\,\partial z}+2\frac{\partial^2\chi_3}{\partial x\,\partial y}\right\}$$

$$=\frac{3m-n}{2n}\left[(\nabla^2(\theta_1+\theta_2+\theta_3)-2\left\{\frac{\partial^2\theta_1}{\partial x^2}+\frac{^2\theta}{\partial y^2}+\frac{\partial^2\theta_3}{\partial z^2}+2\frac{\partial^2\chi_1}{\partial y\,\partial z}+2\frac{\partial^2\chi_2}{\partial x\,\partial z}+2\frac{\partial^2\chi_3}{\partial x\,\partial y}\right\}\right];$$

and finally,

$$\rho\ddot{\phi}-(m+n)\nabla^2\phi+2m\left\{\frac{\partial^2\chi_1}{\partial x^2}+\frac{\partial^2\chi_2}{\partial y^2}+\frac{\partial^2\chi_3}{\partial z^2}+2\frac{\partial^2\psi_1}{\partial y\,\partial z}+2\frac{\partial^2\psi_2}{\partial x\,\partial z}+2\frac{\partial^2\psi_3}{\partial x\,\partial y}\right\}=0. \qquad (9)$$

These are the modified form of (5) and (7).

5. Inclusion of bodily forces, with particular reference to gravity on the surface of the Earth, and to centrifugal forces.

In any case we shall have to consider the alteration made in equations (1), and their solution in (2) by the introduction

on the left-hand side of (1) of terms representing the components of the force per unit volume.

These components can always be represented in a form similar to that given for the displacement in (3), but I shall suppose that the force per unit volume can be represented by the simpler forms

$$\rho\frac{\partial(\mathrm{F}+f_1)}{\partial x},\quad \rho\frac{\partial(\mathrm{F}+f_2)}{\partial y},\quad \rho\frac{\partial(\mathrm{F}+f_3)}{\partial z},$$

and I shall suppose that we have selected the axes and that the ψ's are null. Then in (2) we must replace P by $\mathrm{P}+\rho(\mathrm{F}+f_1)$, Q by $\mathrm{Q}+\rho(\mathrm{F}+f_2)$, R by $\mathrm{R}+\rho(\mathrm{F}+f_3)$, leaving S, T, U unchanged.

In the values found from the stress-strain relations there are no such changes to be made.

We shall thus obtain from (6)

$$\nabla^2\theta_1-\rho f_1=\nabla^2\theta_2-\rho f_2=\nabla^2\theta_3-\rho f_3$$

and

$$2n\{3\rho\mathrm{F}+\rho(f_1+f_2+f_3)\}+(3m+n)\nabla^2(\theta_1+\theta_2+\theta_3)=6m\left\{\frac{\partial^2\theta_1}{\partial x^2}+\frac{\partial^2\theta_2}{\partial y^2}+\frac{\partial^2\theta_3}{\partial z^2}\right\}. \quad (10)$$

If we write

$$\theta_1=\phi+\chi_1+\phi'+\chi_1',\quad \theta_2=\phi+\chi_2+\phi'+\chi_2',$$
$$\theta_3=\phi+\chi_3+\phi'+\chi_3',$$

where ϕ, χ_1, χ_2, χ_3 have the values in (5) and (7) and may be regarded as Complementary Functions, then we remain with

$$\nabla^2\chi_1'=\rho f_1,\quad \nabla^2\chi_2'=\rho f_2,\quad \nabla^2\chi_3'=\rho f_3$$

and

$$2n\rho\mathrm{F}+(m+n)\rho(f_1+f_2+f_3)+(m+n)\nabla^2\phi'=6m\left(\frac{\partial^2\chi_1'}{\partial x}+\frac{\partial^2\chi_2}{\partial y^2}+\frac{\partial^2\chi_3}{\partial z^2}\right), \quad (11)$$

which may be regarded as giving the Particular Integral corresponding to the particular force acting.

There are not many cases of interest. In the case of gravity on the surface of the Earth, as under natural forces generally, we have

$$f_1=f_2=f_3=0 \quad \text{and} \quad (m+n)\nabla^2\phi'+2n\rho\mathrm{F}=0.$$

If we suppose $(-\lambda, -\mu, -\nu)$ to be the direction-cosines of the attraction of gravitation,

$$\mathrm{F}=-g(\lambda x+\mu y+\nu z)$$

and

$$\phi'=\frac{ng\rho}{3(m+n)}(\lambda x^3+\mu y^3+\nu z^3);$$

and if we write P′, Q′, R′ for the Particular Integral portion of P, Q, R—*i. e.* the terms which depend explicitly on g—we find

$$P' = g\rho\lambda x + \frac{m-n}{m+n} g\rho(\mu y + \nu z), \quad . \quad . \quad . \quad (12)$$

with similar values for Q′ and R′, the Complementary Function part of P, Q, and R and the values of S, T, U being those given in the earlier part of this paper.

The other case which I propose to consider is that of a body moving with angular velocities ω_x, ω_y, ω_z about the axes of coordinates which must be axes fixed in the body. It is implied either that the question is purely kinematical, or that a problem in Rigid Dynamics has been previously solved.

The expressions for the acceleration of a point in the body are well known, and give for the effect of the reversed effective forces

$$F = \tfrac{1}{2}\{(\omega_y^2+\omega_z^2)x^2 + (\omega_x^2+\omega_z^2)y^2 + (\omega_x^2+\omega_y^2)z^2 - 2\omega_x\omega_y xy - 2\omega_y\omega_z yz - 2\omega_z\omega_x zx\},$$

with

$$f_1 = x(y\dot{\omega}_z - z\dot{\omega}_y), \quad f_2 = y(z\dot{\omega}_x - x\dot{\omega}_z), \quad f_3 = z(x\dot{\omega}_y - y\dot{\omega}_x).$$

The form of the forces f_1, f_2, f_3 indicates that they will cause no strain in the body, and consequently cause no stress. If we proceed to find the effect which they have on the values of the stresses, they will be seen to disappear from the stress-equations. I shall therefore omit them for this purpose, and treat χ_1', χ_2', χ_3' as null.

We then find

$$\phi' = -\frac{n\rho}{12(m+n)}\{(\omega_y^2+\omega_z^2)x^4 + (\omega_z^2+\omega_x^2)y^4 + (\omega_x^2+\omega_y^2)z^4 - 2\omega_y\omega_z yz(y^2+z^2) - 2\omega_z\omega_x zx(z^2+x^2) - 2\omega_x\omega_y xy(x^2+y^2)\}$$

and

$$P' = -F - f_1 - \left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi', \text{ etc.},$$

$$S' = \frac{\partial^2\phi'}{\partial y\,\partial z}, \text{ etc.}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

with Complementary Function terms as before.

These give no solution of any specific question. They only give a skeleton of the general form which a solution will take.

XXI. *Note on Damped Vibrations.*

By H. S. Rowell *.

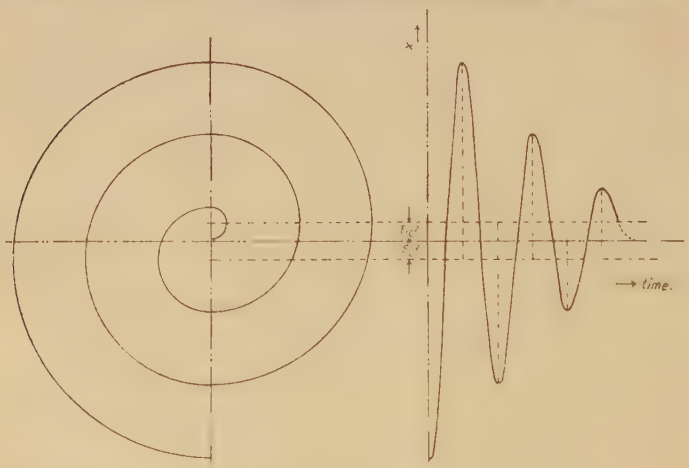
IT is well known that the space time curve for free undamped vibrations may be derived from the projection of a rotating vector, the end of which describes a circle, and it is fairly well known that for vibrations which are resisted by fluid friction proportional to the velocity, the space time curve may be projected (as remarked by P. G. Tait) from a rotating vector, the end of which describes an equiangular or logarithmic spiral.

The vibration of bodies when resisted by a constant frictional force—say solid friction—is of great importance in practical work and does not appear to have been adequately treated. The results obtainable are, moreover, in themselves of much interest.

If F is the constant force of friction the equation of motion is

$$m\ddot{x} + c^2 x \pm \mathrm{F} = 0,$$

wherein the sign of F depends on the direction of motion.



Substituting $x = \mathrm{X} \mp \mathrm{F}/c^2$, we have

$$\mathrm{X} = x \pm \frac{\mathrm{F}}{c^2} = \mathrm{A}\cos\left(\frac{ct}{\sqrt{m}} + \epsilon\right),$$

which gives a series of harmonic vibrations about alternating centres distant F/c^2 from the equilibrium position when

* Communicated by the Author.

friction is absent. The motion can be obtained by projection from a spiral which is composed of semicircles as shown in the diagram.

It will be seen that the amplitudes are in arithmetical progression and the difference for a complete period is $4F/c^2$, which may be called the arithmetic decrement. There is little purpose in using the ordinary definition of decrement, but it may be remarked that on this definition (*i. e.* ratio of successive amplitudes) the decrement ranges from unity for infinite amplitudes to infinity for zero amplitudes.

The spiral curve described here does not appear to have been used before in scientific work, and it might be conveniently called the arithmetic spiral or the spiral of semicircles.

XXII. *Notices respecting New Books.*

Weather Prediction by Numerical Process. By Lewis F. Richardson, B.A., F.Inst.P. 4to, pp. xii+236. Cambridge University Press. 30*s.* net.

THE usual method employed in weather forecasting is a development of that of Abercromby. Distributions of pressure are classified according to standard types, and the variation on any occasion is predicted according to the behaviour of the atmosphere on previous occasions when conditions of the same type occurred. The method is therefore one of sampling inference in which the information utilized is all of one kind. Mr. Richardson believes that other information is relevant to the behaviour of the atmosphere; and in this book he shows how to make use of the known results expressed in the hydrodynamical equations of motion and the equations of emission, transference, and absorption of heat and water. The method adopted is to work with equations each containing only one partial differential coefficient with regard to the time, so that this can be determined by means of the equation when the other quantities involved are known; they include, of course, partial derivatives with regard to the position on the map and the height. These are to be found by observation at stations distributed according to a regular pattern, and the rate of change of each meteorological element at each station is to be calculated from them. Complications arise from the facts that the observations must be made at finite intervals both of time and of position, but these are allowed for. The stations required are more numerous than those at present in operation, and observations should be made every three hours to obtain the best results. Observations of upper-air winds and temperatures are required.

The method is one that appeals strongly to the mathematical physicist. It is necessarily laborious in its present form, and probably could not be worked with sufficient speed to make it a practical method of forecasting; but when forecasters have acquired experience in its use, they will probably find that a sufficient number of the quantities allowed for are comparatively small to make it possible to expedite the calculation considerably without great sacrifice of accuracy.

The value of the work is not confined to the application to forecasting, though the possibility of predicting the disturbing occasions when cyclones cause merriment in the daily press by moving in the wrong direction makes this the feature of most general interest. Its discussion of the physical properties of the atmosphere is so thorough that it constitutes a text-book of the subject. Copious references to original literature are given, and any meteorologist requiring serious information on any topic will do well to look first in this book. The section on evaporation suggests that the only limitation on the evaporation from vegetation is imposed by the difficulty of passing along the stomata tubes; this is not always true even for an isolated leaf, and is certainly wrong for a carpet of grass, on account of the obstruction offered by the vapour from one stoma to evaporation from another. The numerical data actually given, however, eliminate this source of error.

Concerning the printing and style of the book, it is only necessary to say that it is published by the Cambridge University Press. The index is good. H. J.

XXIII. *Proceedings of Learned Societies.*

GEOLOGICAL SOCIETY.

[Continued from vol. xliii. p. 1136.]

February 1st, 1922.—Mr. R. D. Oldham, F.R.S., President, in the Chair.

Mr. Cyril Edward Nowill Bromehead, B.A., F.G.S., delivered a lecture on the Influence of Geology on the History of London.

The 6-inch Geological Survey maps constructed by the Lecturer were exhibited, and some of the new features pointed out. The small streams now 'buried' are indicated on the maps, and the historical research involved in tracing them led to an appreciation of the connexion between the geology and topography on the one hand, and the original settlement and gradual growth of London on the other. The reasons for the first selection of the site have been dealt with by several writers: below London the wide alluvial marshes formed an impassable obstacle; traffic from the

Continent came by the ports of Kent, and, if destined for the north or east of Britain, sought the lowest possible crossing of the Thames. This was near old London Bridge, where the low-level gravel on the south and the Middle Terrace deposits on the north approached close to the river-bank. A settlement was obviously required here, and the northern side was chosen as the higher ground. The gravels provided a dry healthy soil and an easily accessible water-supply; they crowned twin hills separated by the deep valley of the Walbrook, bounded on the east by the low ground near the Tower and the Lea with its marshes, and on the west by the steep descent to the Fleet; the site was, therefore, easily defensible. The river-face of the hills was naturally more abrupt than it is now, owing to the reclamation of ground from the river; the most ancient embankment lay 60 feet north of the northern side of Thames Street.

The first definite evidence of a permanent settlement was the reference in Tacitus. The early Roman encampment lay east of the Walbrook, and the brickearth on the west around St. Paul's was worked. Later the city expanded, until the St. Paul's hill was included, the wall being built in the second half of the 4th century. The great Roman road from Kent (Watling Street) avoided London, and utilized the next ford upstream—at Westminster—on its way to Verulamium and the north-west. The earliest Westminster was a Roman settlement beside the ford, built on a small island of gravel and sand between two mouths of the Tyburn. This settlement could not grow, as did London, since the area of the island, known to the Saxons as Thorney, was small. The road from London to the west joined the St. Alban's road at Hyde Park Corner, running along the 'Strand,' where the gravel came close to the river; a spring thrown out from this gravel by the London Clay was utilized for the Roman Bath in Strand Lane.

Throughout Mediæval times London was practically confined to the walled city, a defensible position being essential. The forests of the London-Clay belt on the north are indicated in Domesday Book and referred to by several writers, notably Fitzstephen, whose Chronicle also mentions many of the springs and wells and the marsh of Moorfields, produced largely by the damming of the Walbrook by the Wall. The same writer mentions that London and Westminster are 'connected by a suburb.' This was along the 'Strand,' and consisted first of great noblemen's houses facing the river and a row of cottages along the north side of the road; this link grew northwards, at first slowly, but in the second half of the 17th century with great rapidity. By the end of that period the whole of the area covered by the Middle-Terrace Gravel was built over, but the northern margin of the gravel was also that of the town for 100 years, the London-Clay belt remaining unoccupied.

The reason for this arrested development was that the gravel

provided the water-supply. In early days the City was dependent on many wells sunk through the gravel, some of which were famous, such as Clerkenwell, Holywell, and St. Clement's. In the same way the outlying hamlets (for instance, Putney, Roehampton, Clapham, Brixton, Ealing, Acton, Paddington, Kensington, Islington, etc.) started on the gravel, but later outgrew it, as pointed out by Prestwich in his Presidential Address of 1873. In the City the supply soon became inadequate, or as Stow says 'decayed,' and sundry means were adopted to supplement it. The conduit system, bringing water in pipes from distant springs, began in 1236; London-Bridge Waterworks pumped water from the Thames by water-wheels from 1582 to 1817; the New River was constructed in 1613, and is still in use. It was not until the 19th century that steam-pumps and iron pipes made it possible for the clay area to be occupied, thus linking together the various hamlets that are now the Metropolitan Boroughs.

Some of the ways in which Geology affects London to-day were briefly indicated, and the lecture was illustrated by a number of lantern-slides, reproduced mainly from old maps and prints.

XXIV. *Intelligence and Miscellaneous Articles.*

YOUNG'S MODULUS AND POISSON'S RATIO FOR SPRUCE.

To the Editors of the Philosophical Magazine.

DEAR SIRS,—

IN my recent paper in the Philosophical Magazine for May 1922 there is an error on page 877. It is there stated that

$$\frac{\sigma_{yz}}{E_y}=\frac{\sigma_{zy}}{E_z},\quad \frac{\sigma_{zx}}{E_z}=\frac{\sigma_{xz}}{E_x},\quad \frac{\sigma_{xy}}{E_x}=\frac{\sigma_{yx}}{E_y}$$

are equal respectively to $700\,S_y$, $700\,S_x$ and $\frac{S_y}{S_x}$.

This should read

$$\frac{E_z}{E_y}=\frac{\sigma_{zy}}{\sigma_{yz}},\quad \frac{E_z}{E_x}=\frac{\sigma_{zx}}{\sigma_{xz}},\quad \frac{E_x}{E_y}=\frac{\sigma_{xy}}{\sigma_{yx}}$$

are equal respectively to $700\,S_y$, $700\,S_x$ and $\frac{S_y}{S_x}$.

The error becomes evident on reading the paper, but I very much regret that it has crept in.

Yours faithfully,
H. CARRINGTON.

The College of Technology,
Manchester.
May 24th, 1922.